

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

MAESTRÍA EN ESTADÍSTICA APLICADA

DETERMINANTES DE LA TEMPORALIDAD EN EL MERCADO LABORAL
ECUATORIANO

**TESIS DE GRADO PREVIO A LA OBTENCIÓN DEL TÍTULO DE MAGISTER
EN ESTADÍSTICA APLICADA**

MERY YANNIRA CHÁVEZ CHAMORRO
yaniris.mericita@gmail.com

Director: PAÚL LEONARDO MEDINA VÁSQUEZ, PHD
pmedinavz@gmail.com

Codirector: LUIS ALCIDES HORNA HUARACA, PHD
luis.horna@epn.edu.ec

QUITO, 2015

DECLARACIÓN

Yo, MERY YANNIRA CHÁVEZ CHAMORRO, declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.

Mery Yannira Chávez Chamorro

CERTIFICACIÓN

Certificamos que el presente trabajo fue desarrollado por MERY YANNIRA CHÁVEZ CHAMORRO, bajo nuestra supervisión.

PAÚL LEONARDO MEDINA VÁSQUEZ, PhD
Director de la Tesis

Luis Alcides Horna Huaraca, PhD
Codirector de la Tesis

Índice general

Resumen	xI
Abstract	xIII
1. Introducción	1
1.1. Problemática	1
1.2. Objetivos del proyecto	3
1.3. Organización del proyecto	3
2. Marco Teórico	5
2.1. Regresión Logística, RL	5
2.1.1. Especificación del modelo	5
2.1.2. Estimación de los coeficientes del modelo y de sus errores estándar	7
2.2. Máquina de vectores soporte MVS	10
2.2.1. Introducción	10
2.2.2. Máquina de vectores soporte para clasificación	10
2.2.3. Problemas con las MVS	13
3. Aplicación	15
3.1. Introducción	16
3.2. Metodología	19
3.3. Modelo	20

3.3.1. Datos	21
3.3.2. Resultados del Modelo	28
3.3.3. Análisis por Rama de Actividad	35
3.4. Conclusiones	47
4. Actualización: Análisis de Resultados	49
4.1. Modelo Logit	49
4.1.1. Comparación con nuevos individuos	54
4.2. Modelo de Máquinas de Vectores Soporte (MVS)	57
4.2.1. Estimación de la MVS	58
5. Conclusiones y Recomendaciones	71
5.1. Conclusiones	71
5.2. Recomendaciones	73
Referencias	74
Anexos	79
A. Análisis Descriptivo	79
B. Validación Cruzada usando K grupos (K-fold cross validation)	84
C. Referencia de Publicaciones	85

Índice de figuras

3.1. Distribución porcentual de individuos por Rama de Actividad.	36
4.1. Esquema de estimación de la MVS.	59
C.1. Publicación Revista de métodos cuantitativos para la economía y la empresa (2012).	85
C.2. Publicación en Proceedings 59th ISI World Statistics Congress (2013).	86

Índice de tablas

3.1.	Frecuencias y Porcentajes para la variable <i>Sexo</i>	22
3.2.	Frecuencias y Porcentajes para la variable <i>Rango de Edad</i>	23
3.3.	Frecuencias y Porcentajes para la variable <i>Nivel de Instrucción</i>	23
3.4.	Frecuencias y Porcentajes para la variable <i>Estado Civil</i>	24
3.5.	Frecuencias y Porcentajes para la variable <i>Ocupación</i>	25
3.6.	Frecuencias y Porcentajes para la variable <i>Antigüedad</i>	25
3.7.	Frecuencias y Porcentajes para la variable <i>Jornada Laboral</i>	25
3.8.	Frecuencias y Porcentajes para la variable <i>Tamaño del establecimiento</i>	26
3.9.	Frecuencias y Porcentajes para la variable <i>Capacitación</i>	26
3.10.	Frecuencias y Porcentajes para la variable <i>Ciudades Autorepresentadas y Regiones</i>	27
3.11.	Resultados de las variables incluidas en el Modelo.	29
3.12.	Tabla de Clasificación.	31
3.13.	Resumen de la Validación del Modelo.	31
3.14.	Escenarios para el Modelo.	34
3.15.	Caracterización de los individuos con contrato temporal en la Rama A.	37
3.16.	Caracterización de los individuos con contrato temporal en la Rama B.	37
3.17.	Caracterización de los individuos con contrato temporal en la Rama C.	38

3.18.	Caracterización de los individuos con contrato temporal en la Rama D.	39
3.19.	Caracterización de los individuos con contrato temporal en la Rama E.	39
3.20.	Caracterización de los individuos con contrato temporal en la Rama F.	40
3.21.	Caracterización de los individuos con contrato temporal en la Rama G.	40
3.22.	Caracterización de los individuos con contrato temporal en la Rama H.	41
3.23.	Caracterización de los individuos con contrato temporal en la Rama I.	41
3.24.	Caracterización de los individuos con contrato temporal en la Rama J.	42
3.25.	Caracterización de los individuos con contrato temporal en la Rama K.	42
3.26.	Caracterización de los individuos con contrato temporal en la Rama L.	43
3.27.	Caracterización de los individuos con contrato temporal en la Rama M.	44
3.28.	Caracterización de los individuos con contrato temporal en la Rama N.	44
3.29.	Caracterización de los individuos con contrato temporal en la Rama O.	45
3.30.	Caracterización de los individuos con contrato temporal en la Rama P.	45
3.31.	Factores de riesgo y protección por Rama de Actividad.	47
4.1.	Resultados de clasificación de los estudios.	50
4.2.	Resultados de los estudios.	52
4.3.	Resultados de los estudios.	53
4.4.	Tabla de clasificación con nuevos individuos del año 2012.	55
4.5.	Tabla de clasificación con nuevos individuos del año 2013.	55

4.6.	Tabla de clasificación con nuevos individuos del año 2014.	56
4.7.	Elección de parámetros C y γ	60
4.8.	Parámetros C y γ para modelos agregados.	60
4.9.	Tabla de clasificación del estudio jun07jun10 con kernel Radial $C = 100$ y $\gamma = 0,001$	61
4.10.	Tabla de clasificación del periodo jun07jun10 con kernel Lineal. . .	61
4.11.	Tabla de clasificación del periodo jun07jun10 con kernel Polinomial $C = 100$ y $\gamma = 0,001$	61
4.12.	Tabla de clasificación del periodo jun07jun10 con kernel Sigmoidal. . .	62
4.13.	Tabla de clasificación del periodo jun07jun12 con kernel Radial $C = 100$ y $\gamma = 0,001$	62
4.14.	Tabla de clasificación del periodo jun07jun12 con kernel Lineal. . .	63
4.15.	Tabla de clasificación del periodo jun07jun12 con kernel Polinomial $C = 100$ y $\gamma = 0,001$	63
4.16.	Tabla de clasificación del periodo jun07jun12 con kernel Sigmoidal. . .	64
4.17.	Tabla de clasificación de nuevos individuos diciembre 2012.	65
4.18.	Tabla de clasificación de nuevos individuos kernel Lineal.	65
4.19.	Tabla de clasificación con nuevos individuos con kernel Polinomial. . .	65
4.20.	Tabla de clasificación con nuevos individuos con kernel Sigmoidal. . .	66
4.21.	Tabla de clasificación con nuevos individuos con kernel Radial. . . .	66
4.22.	Tabla de clasificación con nuevos individuos con kernel Lineal. . . .	67
4.23.	Tabla de clasificación con nuevos individuos con kernel Polinomial. . .	67
4.24.	Tabla de clasificación con nuevos individuos con kernel Sigmoidal. . .	68
4.25.	Tabla de clasificación con nuevos individuos con kernel Radial. . . .	68
4.26.	Tabla de clasificación con nuevos individuos con kernel Lineal. . . .	69
4.27.	Tabla de clasificación con nuevos individuos con kernel Polinomial. . .	69
4.28.	Tabla de clasificación con nuevos individuos con kernel Sigmoidal. . .	70
A.1.	Distribución por tipo de contrato de la variable Sexo.	80

A.2.	Distribución por tipo de contrato de la variable <i>Edad</i>	80
A.3.	Distribución por tipo de contrato de la variable <i>Nivel de Instrucción</i>	80
A.4.	Distribución por tipo de contrato de la variable <i>Estado Civil</i>	80
A.5.	Distribución por tipo de contrato de la variable <i>Ocupación</i>	81
A.6.	Distribución por tipo de contrato de la variable <i>Antigüedad</i>	81
A.7.	Distribución por tipo de contrato de la variable <i>Jornada Laboral</i>	82
A.8.	Distribución por tipo de contrato de la variable <i>Tamaño del Establecimiento</i>	82
A.9.	Distribución por tipo de contrato de la variable <i>Capacitación</i>	82
A.10.	Distribución por tipo de contrato de la variable <i>Ciudades Autorepresentadas y Regiones</i>	83

Resumen

Las diferencias que existen dentro del mercado laboral ecuatoriano, en las diferentes ramas de actividad, evidencian los niveles de formación que los trabajadores deben tener para evitar efectos negativos. Por tal motivo, en este estudio se analiza el papel que las características personales, laborales y geográficas desempeñan en la probabilidad de obtener un contrato temporal frente a uno indefinido. El análisis es realizado por rama de actividad, para determinar qué características posibilitarían la existencia del contrato temporal en cada una de ellas. Para lograr este objetivo, se desarrollan modelos de regresión logística y luego, en segundo lugar, a manera de contraste o validación se desarrollan modelos denominados de aprendizaje; en particular, la técnica denominada máquina de vectores soporte. En ambos casos la información utilizada se basó utilizando los datos de la Encuesta de Empleo, Desempleo y Subempleo, elaborada por el Instituto Nacional de Estadística y Censos (INEC).

A partir de lo indicado, en este estudio se presentarán los fundamentos matemáticos de la regresión logística y de la máquina de vectores soporte. Para ello, se prestará especial atención a los aspectos más relevantes y necesarios para el desarrollo e implementación de las técnicas, enfocándose, principalmente, en un problema de clasificación binario. Bajo el contexto de la técnica de regresión logística se detectan las diferencias que existen dentro del mercado laboral ecuatoriano. En relación a esto, se presenta, de manera particular, en el capítulo 3, la investigación publicada en la “Revista de Métodos Cuantitativos para la Economía y la Empresa”; revista indexada en Scopus, DOAJ, RePEc, ECONLIT, etc., con clasificación Q2 (<http://scimagojr.com/journalsearch.php?q=10400153309&tip=sid>). Cabe resaltar que la autoría y coautoría de esta publicación corresponden a la postulante y al director del presente proyecto, res-

pectivamente. A continuación, se desarrolla la técnica de máquina de vectores soporte con el fin de contrastar los valores obtenidos, para poder mostrar las posibles ventajas y/o desventajas de utilizar un método de aprendizaje para clasificación de individuos, procedimiento muy utilizado en los actuales momentos. Finalmente, se exponen las conclusiones del estudio realizado y se discuten las posibles líneas futuras de trabajo.

Abstract

The differences of diverse activity branches inside the Ecuadorian labor market show the training levels that workers must have to avoid negative effects. Therefore, this study analyzes the role played by personal, labor, and geographical characteristics in the probability of obtaining a temporary employment contract against an undefined one. The analysis is performed by activity branch in order to determine which are the characteristics that would make possible the existence of temporary employment contracts in each of them. In order to achieve this purpose, two stages are developed. First of all, we use a logistic regression model. Then in a second place, we use learning models as a robust technique to check the results. Particularly the technique called support vector machine. In both cases the information used is based on the Survey of Employment, Unemployment and Underemployment, from the National Institute of Statistics and Censuses (INEC).

From the above, the mathematical foundations of logistic regression and support vector machine are presented in this paper. It is given particular attention to the most relevant and necessary aspects for the development and implementation of the techniques, it is mainly focused on a binary classification problem. In the context of the logistic regression technique, there are detected differences that exist within the Ecuadorian labor market. So, Chapter 3 presents in a particular way the entirely research published in the "Journal of Quantitative Methods for Economics and Business". This journal is indexed in Scopus, DOAJ, RePEc, ECONLIT. It is, rated Q2 (<http://scimagojr.com/journalsearch.php?q=10400153309&tip=sid>). However, the authorship of this publication corresponds to the candidate and the advisor of this current project as author and co-author, respectively. Then, the technique of support vector machine are develo-

ped in order to contrast the obtained results. In addition, to present the possible advantages and/or disadvantages of using a learning method for individuals classification, it is a widely used procedure currently. Finally, conclusions of this study are highlighted and future research directions are discussed.

Capítulo 1

Introducción

Este primer capítulo ofrece las definiciones y razonamientos que han motivado la perspectiva que enmarca el presente proyecto. Este se convierte, por lo tanto, en un punto de referencia conveniente para entender los resultados desarrollados con posterioridad. Inicialmente, se define y sitúa el marco de investigación general sobre el que se desarrolla el problema a estudiar. De esta manera, se permite entender las necesidades y la importancia del fenómeno. Luego, se explica el porqué la elección de un modelo de clasificación que busca la construcción, a partir de una muestra, de una función para discriminar nuevos elementos.

1.1. Problemática

Los contratos temporales son a menudo considerados como un importante componente de la flexibilidad del mercado laboral. Quienes trabajan bajo esta modalidad pueden ser despedidos sin incurrir en pago o restricciones impuestas por la legislación sobre derechos de los empleados. Esto podría explicar el crecimiento de esta modalidad laboral en algunos países; como en el caso de Colombia, donde los datos de la encuesta manufacturera anual indican que la proporción de empleo temporal creció de 15 % en 1995 a 48 % en 2007 [7].

En el caso concreto de Ecuador, el crecimiento dentro del mercado laboral ecuatoriano se ha basado en la precarización del trabajador. Esto se ha realizado, olvidando que la productividad de los empleados depende no solo de su aptitud

y de la inversión que se realiza en ello, tanto dentro como fuera del puesto de trabajo, sino asimismo, de su motivación y de la intensidad de su esfuerzo [14]. Actualmente, la estabilidad se ha convertido en una incertidumbre para la mayoría de las personas y resulta cada vez más complejo determinar las causas que posibilitan el acceso a un trabajo estable. Dentro de las teorías económicas que nos ayudan a evidenciar factores que influyen a disminuir la inestabilidad, se puede destacar la teoría del capital humano que señala que “la inversión en la formación de los trabajadores hace que este se convierta en un factor para disminuir el desempleo” [30].

Dado que este “problema” es complejo, pues depende de muchos factores (sexo, educación, edad, región, etc) se construirá un modelo econométrico que permita analizar y estudiar el fenómeno. En primera instancia se recurrirá a una técnica clásica “Regresión Logística (RL)”, propuesta y estudiada por Cornfiel en 1962, Day & Kerridge en 1967, Hosmer & Lemeshow en 1989 y Cox en 1996. Esta por sus características (versatilidad, desarrollo computacional) y su uso es una de las técnicas que más se encuentran en la literatura para dar solución a los problemas de clasificación [40].

Luego de esto se estudiará la validez y robustez del modelo clásico obtenido utilizando la técnica contemporánea “Máquina de Vectores Soporte (MVS)”, propuesta y estudiada por Cortes & Vapnik en 1995; debido a que, en los últimos años, esta técnica ha demostrado excelentes resultados en una gran variedad de problemas de clasificación y de regresión en diferentes áreas (Tripathi en 2006, Crisler en 2008, Hongdong en 2009) [9, 40].

Se recurrirá a las dos técnicas antes señaladas porque son herramientas que sirven para solucionar problemas de clasificación. Como este proyecto se enfocará en un problema de clasificación binomial, era necesario encontrar modelos adecuados para su estudio. Los métodos de clasificación buscan la construcción de una función que, a partir de una muestra, permita discriminar nuevos elementos, es decir, se busca que esta función pueda pronosticar de una manera óptima el grupo al que pertenece una nueva observación.

Finalmente cabe indicar que, como característica particular, el enfoque clásico (RL) para abordar un problema de clasificación como el que se ha planteado, en el que se cuenta con una gran cantidad de datos, consiste en asumir que tales

datos están generados por una distribución de probabilidad subyacente que es desconocida y, a partir de ella se diseñará al clasificador. En cambio, el enfoque contemporáneo (MVS) toma la idea básica de diseñar el clasificador directamente desde los datos mediante determinados algoritmos.

1.2. Objetivos del proyecto

Como se ha indicado anteriormente, este proyecto de titulación busca resolver un problema de clasificación para el mercado laboral ecuatoriano; con este propósito, a continuación se definen dos tipos de objetivos: técnico y econométrico:

- El objetivo técnico es analizar y comparar una técnica clásica, regresión logística (RL), versus una contemporánea, máquina de vectores soporte (MVS), con la intención de averiguar sus alcances y limitaciones.
- El objetivo econométrico es investigar las principales características económicas que influyen para la contratación temporal y determinar en qué medida afectan al momento de conseguir un empleo.

1.3. Organización del proyecto

Una vez efectuada la introducción del trabajo, se lo presenta con detalle en los siguientes capítulos. La organización del proyecto es la siguiente: en el capítulo 2 se presentan los fundamentos matemáticos de la regresión logística (RL) y de la máquina de vectores soporte (MVS), enfocándose en un problema de clasificación binario. En el capítulo 3 se presenta de forma íntegra el trabajo realizado para detectar las diferencias que existen dentro del mercado laboral ecuatoriano el cual se encuentra publicado en la Revista de Métodos Cuantitativos para la Economía y la Empresa. Cabe indicar que la publicación señalada es original, inédita y que no ha servido para ningún concurso o postulación pública o privada, salvo como origen y motivación para este proyecto de titulación. A continuación, en el capítulo 4 se realiza un análisis de los resultados obtenidos al comparar dos técnicas estadísticas utilizadas para la clasificación de indivi-

duos, como son la regresión logística y la máquina de vectores soporte. Dentro del capítulo 5 se exponen las conclusiones del estudio realizado y se discuten las posibles futuras líneas de trabajo. Por último, se incluyen los anexos en los que se presenta un análisis descriptivo de los datos utilizados, una explicación referente a la validación cruzada usando K grupos y la primera página de los trabajos publicados.

Capítulo 2

Marco Teórico

En este capítulo, con el fin de lograr los objetivos planteados en el presente proyecto de titulación, se revisarán los aspectos teóricos de la Regresión Logística (RL) y de las Máquinas de Vectores Soporte (MVS). Estas dos técnicas estadísticas se utilizarán para la estimación de parámetros y la comparación de resultados.

2.1. Regresión Logística, RL

Al ser la regresión logística una técnica ampliamente utilizada, estudiada y difundida, la fundamentación teórica que se expondrá a continuación se basará en textos referenciales, como [15], [34], [35] [36] y, por tal motivo, muchos de los conceptos serán copia textual de los mismos, indicando ciertamente su respectiva autoría.

2.1.1. Especificación del modelo

“Los modelos de regresión logística son modelos estadísticos en los que se desea conocer la relación entre: una variable dependiente cualitativa, dicotómica (regresión logística binaria o binomial) o con más de dos valores (regresión logística multinomial)” [36]. En esta investigación, de manera particular, se tomará como punto central el desarrollo y aplicación de la *regresión logística binaria*.

Se parte del caso de una variable dependiente binaria, Y , la cual viene expli-

cada por un conjunto de variables $X = (x_1, x_2, \dots, x_k)$, la idea consiste en utilizar un modelo de la forma

$$Y = f(x_1, x_2, \dots, x_k) + \mu, \quad (2.1)$$

donde f es una función real que depende de las variables x_1, x_2, \dots, x_k [35]. En particular, el modelo logístico considera f de la forma

$$f(z) = \frac{\exp(z)}{1 + \exp(z)}, \quad (2.2)$$

con $z = \beta'X$. Ahora, si se estima $p_i = f(\beta'X_i)$ como la probabilidad de que la variable respuesta $Y = y_i$, tome en la población el valor de 1 cuando $X = X_i$, [$p_i = P(Y = y_i = 1|X_i)$], tendremos que

$$p_i = \frac{1}{1 + \exp(-\beta'X_i)}. \quad (2.3)$$

Lo cual nos indica que “el modelo logit asume una relación suave entre el vector de variables explicativas y la probabilidad de ocurrencia del suceso” [34]. Además, cuando “ $(\beta'X_i)$ es muy grande y positivo la probabilidad tiende a 1 y cuando este término es muy grande y negativo la probabilidad tiende a cero. Si $\beta'X_i = 0$, la probabilidad es 0.5. La respuesta es simétrica, ya que

$$\begin{aligned} 1 - p_i &= \frac{\exp(-\beta'X_i)}{1 + \exp(-\beta'X_i)} \\ &= \frac{1}{1 + \exp(\beta'X_i)}. \end{aligned} \quad (2.4)$$

Dividiendo estas dos expresiones, (2.3) y (2.4), y tomando logaritmos tendremos que

$$\begin{aligned} g_i &= \ln\left(\frac{p_i}{1 - p_i}\right) \\ &= \beta'X_i \end{aligned} \quad (2.5)$$

que es otra forma conveniente de escribir el modelo logit” [34].

Por otra parte, se debe tener presente que “en este tipo de modelos no resulta posible interpretar directamente las estimaciones de los parámetros β , ya que son modelos no lineales. Lo que haremos en la práctica es fijarnos en el signo de los estimadores. Si el estimador es positivo, significará que incrementos en la variable asociada causan incrementos en $P(Y = 1)$ (aunque desconocemos

la magnitud de los mismos). Por el contrario, si el estimador muestra un signo negativo, ello supondrá que incrementos en la variable asociada causarán disminuciones en $P(Y = 1)$ [36].

Para el análisis del modelo logit se suele utilizar dos conceptos con el fin de profundizar la interpretación de los estimadores.

- Se llama *odds* al siguiente cociente de probabilidades

$$\begin{aligned} O_i &= \exp(g_i) \\ &= \frac{p_i}{1 - p_i} \\ &= \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k) \end{aligned} \quad (2.6)$$

- Se llama *odds-ratio* al cociente entre dos *odds* asociados (el obtenido tras realizar el incremento y el anterior al mismo). En particular, si suponemos que " $x_{i1} = x_{j1} + 1$ y $x_{ih} = x_{jh}$ para $h \geq 2$; entonces, el cociente de los ratios (*odds-ratio*) para estas dos observaciones es

$$\frac{O_i}{O_j} = \exp(\beta_1). \quad (2.7)$$

y el parámetro $\exp(\beta_1)$ indica cuánto se modifica este ratio entre estos dos elementos cuando la variable x_1 aumenta una unidad. Aquí, si la variable x_1 es cualitativa y toma los valores 0 y 1, indica el cambio entre los ratios de disparidad al pasar de un grupo a otro" [34].

2.1.2. Estimación de los coeficientes del modelo y de sus errores estándar

Para la estimación de los coeficientes del modelo y de sus errores estándar se recurre al cálculo de estimaciones de máxima verosimilitud, es decir, estimaciones que maximicen la probabilidad de obtener los valores de la variable dependiente Y proporcionados por los datos de la muestra [36]. Por ejemplo, "supongamos que se dispone de una muestra de n elementos caracterizados por un vector de variables X y un valor (cero o uno) de la variables respuesta

binaria. La función de probabilidades de una respuesta, y_i cualquiera es:

$$P(y_i) = p_i^{y_i}(1 - p_i)^{1-y_i}, \quad y_i = 0, 1 \quad (2.8)$$

y para toda la muestra, suponiendo independencia entre las observaciones:

$$P(y_1, \dots, y_n) = \prod_{i=1}^n p_i^{y_i}(1 - p_i)^{1-y_i}. \quad (2.9)$$

Tomando logaritmos

$$\log(P(\mathbf{y})) = \sum_{i=1}^n [y_i \log p_i + (1 - y_i) \log(1 - p_i)]. \quad (2.10)$$

Esta ecuación indica que maximizar la verosimilitud equivale a minimizar la suma de n términos que miden el ajuste del modelo a los datos" [34].

Estas estimaciones no son de cálculo directo, como ocurre en el caso de las estimaciones de los coeficientes de regresión, por el método de los mínimos cuadrados. Para el cálculo de estimaciones máximo – verosímiles se recurre a métodos iterativos, como el método de Newton – Raphson [36]. Para más detalles ver [34]. Las estimaciones de los coeficientes de regresión no surgen solamente de estos métodos, sino también de sus errores estándar y de las covarianzas entre las covariables del modelo. El siguiente paso será comprobar la significación estadística de cada uno de los coeficientes de regresión en el modelo. Para ello se pueden emplear básicamente tres métodos: el estadístico de *Wald*, el estadístico *G* de razón de verosimilitud y la prueba *Score*. El concepto y resumen de cada uno de ellos que se muestra a continuación ha sido tomado de [36].

El estadístico de Wald

Además de contrastar la hipótesis de que un coeficiente aislado es distinto de 0, y sigue una distribución normal de media 0 y varianza 1. El valor para un coeficiente concreto viene dado por el cociente entre el valor del coeficiente y su correspondiente error estándar. La obtención de significación indica que dicho coeficiente es diferente de 0 y merece la pena su conservación en el modelo. En el caso de los modelos con errores estándares grandes, el estadístico de Wald puede proporcionar falsas ausencias de significación es decir, se incrementa el error tipo II. Tampoco es recomendable su uso si se están empleando variables

de diseño.

El estadístico G de razón de verosimilitud

Se trata de ir contrastando cada modelo que surge después de eliminar de forma aislada cada una de las covariables frente al modelo completo. En este caso cada estadístico G sigue una χ^2 con un grado de libertad (no se asume normalidad). La ausencia de significación implica que el modelo sin la covariable no empeora respecto al modelo completo, en otras palabras da igual su presencia o su ausencia, por lo que, según la estrategia de obtención del modelo más reducido (principio de parsimonia), dicha covariable debe ser eliminada del modelo, pues no aporta nada al mismo. Esta prueba no asume ninguna distribución concreta, por esta razón es la más recomendada para estudiar la significación de los coeficientes.

La prueba Score

Su cálculo para el caso de una única variable viene dado por

$$S = \frac{\sum_{i=1}^n x_i(y_i - \bar{y})}{\sqrt{\bar{y}(1 - \bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2}} \quad (2.11)$$

En el caso de múltiples covariables se debe utilizar cálculo matricial. Si bien no requiere un cálculo iterativo (precisamente su rapidez de cálculo sería su aspecto más favorable), tiene dos aspectos en su contra:

1. Se sabe que este estadístico se incrementa conforme aumenta el número de covariables es decir, tiende a dar significación con mayor frecuencia.
2. Este estadístico también asume una distribución normal con media 0 y varianza 1.

Al igual que en los casos anteriores, si alcanza significación indica que la covariable debería permanecer en el modelo. Su uso en algunos paquetes estadísticos ha quedado relegado a la selección de variables en métodos paso a paso (por la mayor rapidez de cálculo).

Cuando la covariable es cualitativa con n categorías (siendo $n > 2$), en el modelo se analizará la significación de cada una de sus $n + 1$ variables ficticias;

así como la significación global de la covariable, comparando la presencia en bloque frente a la ausencia en bloque de sus $n + 1$ covariables ficticias.

2.2. Máquina de vectores soporte MVS

La técnica máquina de vectores soporte fue desarrollada dentro de la teoría de aprendizaje. A partir de una cantidad finita de datos, esta busca una adecuada función que permita una buena generalización entre un particular conjunto de ensayo y la máquina. A continuación se expondrán los conceptos esenciales y las herramientas matemáticas que constituyen la base teórica de la máquina de vectores soporte para clasificación.

2.2.1. Introducción

La teoría de la máquina de vectores soporte fue desarrollada inicialmente por Vapnik y Lerner (1963) [43] para la resolución de problemas de clasificación linealmente separables mediante lo que se denomina hiperplano óptimo de separación. Después, se amplió su ámbito de trabajo a la resolución de problemas de clasificación no separables a través de algoritmos no lineales; propuesta realizada por Boser (1992), Guyon (1993) Cortes y Vapnik (1995) [43]. Posteriormente, Vapnik (1995), Dimola (1996) y Drucker (1997) [43] extendieron su uso para problemas de regresión. La técnica, en su conjunto, se centra en dos conceptos previamente establecidos: i) el uso de los *kernels* o *núcleos* y su interpretación geométrica y ii) la construcción de un hiperplano de separación óptimo en un contexto no paramétrico.

2.2.2. Máquina de vectores soporte para clasificación

Sea un problema genérico de clasificación o de regresión en el que se dispone de un conjunto etiquetado de entrenamiento con n muestras i.i.d.¹, tal que

$$A = \{(x_1, y_1), \dots, (x_n, y_n)\}, \quad (2.12)$$

¹i.i.d: independientes e idénticamente distribuidas

donde $x_i \in \mathbb{R}^d$ e $y_i \in \{-1, +1\}$. A partir de este conjunto, se puede definir un hiperplano de separación $S(x)$, como una función lineal; de forma que

$$\begin{aligned} S(x) &= (w_1x_1 + \dots + w_dx_d) + b \\ &= \langle w, x \rangle + b, \quad w_i, b \in \mathbb{R}. \end{aligned} \quad (2.13)$$

En particular, se cumplirá que

$$S(x) = \langle w, x \rangle + b = \begin{cases} \langle w, x_i \rangle + b \geq 0 & \text{si } y_i = +1 \\ \langle w, x_i \rangle + b \leq 0 & \text{si } y_i = -1 \end{cases}$$

donde $i = 1, \dots, n$. $\langle w, x_i \rangle + b \geq 0$ y $\langle w, x_i \rangle + b \leq 0$, se los conoce como hiperplanos positivo y negativo, respectivamente.

Aquí se debe tener presente que el hiperplano de separación $S(x)$, tal y como está definido, no es único; es decir, existen infinitos hiperplanos de separación, caracterizados por todos los hiperplanos que son capaces de cumplir las restricciones impuestas. Sin embargo, se sabe que la distancia más corta entre un hiperplano positivo y el punto positivo más cercano es igual a

$$\frac{1}{\|w\|}. \quad (2.14)$$

La misma definición se cumple en el caso del hiperplano negativo. Como consecuencia, se puede decir que la distancia entre los hiperplanos positivo y negativo, conocida como *margen*, es igual a

$$\frac{2}{\|w\|}. \quad (2.15)$$

Los puntos que están ubicados a ambos lados del hiperplano óptimo

$$\langle w, x_i \rangle + b = 0 \quad (2.16)$$

y que, por lo tanto, definen el margen, se denominan *vectores soporte*. Como complemento a esto, la búsqueda del hiperplano óptimo estará únicamente direccionada a encontrar los valores w y b que minimizan el funcional $\|w\|$ o maximizan el margen, sujeto a la restricción $y_i(\langle w, x_i \rangle + b) - 1 \geq 0$. En otras

palabras,

$$\text{mín } \frac{1}{2} \|w\|^2$$

$$\text{s.a. } y_i(\langle w, x_i \rangle + b) - 1 \geq 0.$$

El vector w define el hiperplano de decisión en dicho espacio y b representa el sesgo del origen de coordenadas. Esto, en principio, puede ser resuelto con alguna técnica de la teoría de optimización, pues es un problema de programación cuadrática.

Por otra parte, para el caso específico de la clasificación binaria es conveniente definir una función de transformación Φ :

$$\Phi : \mathbb{X} \rightarrow \mathbb{H},$$

tal que, hace corresponder a cada vector de entrada x con un punto en el espacio de características \mathbb{H} . Bajo esta consideración, la idea será construir un hiperplano de separación lineal en este nuevo espacio. De tal forma que la frontera de decisión final obtenida en el espacio de características, se transformará en una frontera de decisión no lineal en el espacio original de entradas [12]. Específicamente, definiremos $f(x)$, tal que

$$\begin{aligned} f(x) &= w^T \Phi(x) + b, \\ &= \langle w, \Phi(x) \rangle + b \end{aligned} \quad (2.17)$$

y, en su forma dual

$$f(x) = \sum_{i=1}^n \alpha_i^* y_i K(x, x_i), \quad (2.18)$$

donde $K(x, x')$ se conoce como función *kernel* o *núcleo*. De manera particular, una función kernel será una función que asigne a cada par de elementos del espacio de entrada $\mathbb{X} \times \mathbb{X}$ un valor real. Este será igual al producto escalar de las imágenes de dichos elementos en el espacio de características \mathbb{H} , es decir,

$$\begin{aligned} K : \mathbb{X} \times \mathbb{X} &\rightarrow \mathbb{R} \\ (x, x') &\mapsto \langle \Phi(x), \Phi(x') \rangle = K(x, x') \\ &= (\Phi_1(x)\Phi_1(x') + \dots + \Phi_d(x)\Phi_d(x')). \end{aligned}$$

Puesto que el objetivo para resolver el problema de clasificación se centraba

en minimizar el funcional $\|w\|$, sujeto a la restricción $y_i(\langle w, x_i \rangle + b) - 1 \geq 0$, bajo la introducción de la función Φ y del mismo kernel $K(x, x')$; el resolver este problema de minimización será equivalente a solucionar el siguiente problema dual

$$\text{máx } \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{(i,j)=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (2.19)$$

$$\text{s.a. : } \sum_{i=1}^n \alpha_i y_i = 0 \quad (2.20)$$

$$0 \leq \alpha_i \leq C. \quad (2.21)$$

Allí, $i = 1, \dots, n$, C es un parámetro de regularización que permite controlar el número de errores de clasificación permitidos en la etapa de aprendizaje. “Actualmente, no existe una forma teórica de encontrar el valor C . Solo existe la heurística de usar un valor grande” [12].

Para resolver este problema dual no es necesario conocer el conjunto de funciones base Φ , ni tampoco, de manera explícita, las coordenadas de los datos (puntos) a clasificar en el espacio de características. Simplemente basta con conocer la forma funcional del kernel correspondiente $K(x, x')$, aun cuando este pudiese estar asociado a un conjunto finito de funciones base. Para más detalles véase [12],[43],[40].

Las funciones kernels más usadas y recomendadas en [33] y [38] son:

- Kernel lineal: $K(x, x') = \langle x, x' \rangle$.
- Kernel polinomial de grado d : $K_d(x, x') = [\gamma \langle x, x' \rangle + \gamma]^d$.
- Kernel gaussiano $K(x, x') = \exp(-\gamma \|x - x'\|^2)$, son: $\gamma > 0$
- Kernel sigmoidal: $K(x, x') = \tanh(\gamma \langle x, x' \rangle + \tau)$

A los parámetros γ, τ, d se les denomina parámetros del kernel o núcleo.

2.2.3. Problemas con las MVS

Los dos problemas más comunes con la MVS son el *overtraining* y el *overfitting*.

- **Overtraining:** ocurre cuando se han aprendido muy bien los datos de entrenamiento pero no se pueden clasificar de manera adecuada ejemplos

nunca antes vistos (datos de verificación); es decir, se trata de una mala generalización del modelo.

- **Overfitting:** se presenta cuando no se ha aprendido muy bien la característica de los datos de entrenamiento, por lo que se hace una mala clasificación.

Se debe tener en cuenta estas consideraciones a la hora de ajustar el modelo ya que de ello depende la exactitud y éxito de la predicción.

Capítulo 3

Aplicación

En este capítulo, se presentará el artículo [14], del cual se deriva esta tesis de maestría, tal como se lo publicó en la Revista de Métodos Cuantitativos para la Economía y la Empresa. Esta revista científica se encuentra indexada tanto en bases multidisciplinarias (International Bibliography of Social Sciences) como en bases especializadas (EconLit, DOAJ) y bases de citas (Scopus).

Esencialmente, se muestra la aplicación de la regresión logística a un problema de la sociedad humana fundamental: el empleo. En resumen, tomando en cuenta distintas características cuantitativas y cualitativas (sexo, edad, formación, etc.) se analiza la probabilidad de que un determinado individuo hombre/mujer pueda acceder a un empleo de forma temporal o indefinida.

Cabe indicar que el artículo señalado no ha sido presentado antes ni forma parte de otra tesis, proyecto de titulación o concurso de merecimientos. El artículo es original y fue desarrollado en el transcurso de la maestría (Ver Anexo C). Su autoría corresponde a la proponente de este proyecto de titulación y a su director.

Determinantes de la temporalidad en el mercado ecuatoriano

Resumen

Las diferencias que existen dentro del mercado laboral ecuatoriano, en las diferentes ramas de actividad, evidencian los niveles de formación que los trabajadores deben tener para evitar efectos negativos. Por tal motivo, en este estudio se analiza el papel que las características personales, laborales y geo-

gráficas desempeñan en la probabilidad de obtener un contrato temporal frente a uno indefinido. El análisis es realizado por rama de actividad, para determinar qué características posibilitarían la existencia del contrato temporal en cada una de ellas. Para lograr este objetivo, se estiman modelos de regresión logística utilizando los datos de la Encuesta de Empleo, Desempleo y Subempleo desde el II trimestre del año 2007 al II trimestre del año 2010, elaborada por el Instituto Nacional de Estadística y Censos (INEC).

Palabras claves: contrato temporal, contrato indefinido, regresión logística, rama de actividad.

Códigos JEL: J23, J24, J41. *MSC2000:* 00A71, 03C98, 93E10.

Abstract

The differences that exists inside the labor ecuadorian market, in the different branches of activity, they demonstrate the training levels that the workers must have to avoid negative effects. However, in this study, the role played by individual, jobs and residence characteristics are analyzed on the probability of having a fixed-term employment versus permanent employment. It is analyzed concretely by branch of activity, to check which are the characteristics that would make possible the existence of the fixed-term employment in each of them. To achieve this aim there are estimated models of logistic regression using the information of the Survey of Employment, Unemployment and Underemployment from 2nd quarter of 2007 to the 2nd quarter of 2010, elaborated by the National Institute of Statistics and Censuses (INEC).

Keywords: fixed-term employment, permanent employment, logistic regression, branch of activity.

3.1. Introducción

El mercado laboral ecuadoriano ha vivido cambios estructurales en las décadas de los 80s, 90s y en el primer lustro del nuevo siglo, cuando se empezó con las propuestas de la flexibilidad laboral, una mayor apertura comercial, las propuestas privatizadoras, la apertura financiera, las reformas de la política cambiaria y la reforma tributaria; fueron los puntales básicos de este cambio [1]. Las

tasas de desempleo han estado en un rango del 8% al 15%, con su máximo en la crisis de 1999 y su mínimo en el año de 1980. El subempleo, ha tenido unas tasas igualmente altas, llegando hasta el 54% en 1991 y el nivel más bajo en 1980, con el 30%. En el período de 1999 a 2006, los niveles de desempleo se estabilizaron a niveles similares a los de la pre-crisis (1994-1998). En este período se tuvo un pico del 17% de desempleo total en febrero del 2000; posteriormente, el desempleo más alto que se tuvo fue del 12,05% en abril del 2004. El mes con menos desempleo fue diciembre del 2002 con el 7,74% [5]. En el período del 2007 al 2010¹ se presenta una mejoría sensible a la registrada en las últimas décadas, manteniéndose en promedio una tasa del 7,5%; la tasa más alta se registra en marzo del 2010 con un 9,1% y la más baja en diciembre del 2010 con un 6,1% [23].

La flexibilidad laboral en el Ecuador empezó desde 1990 con la Ley de Maquilas y la contratación a tiempo parcial [28], proporcionando otras formas de contratación como: contratos por tiempo parcial y contratos eventuales. Dentro del concepto de contrato temporal se distinguen dos grupos. El primero considera a los contratos eventuales, es decir, contratos por obra, a destajo, por horas y por jornal. Estos tipos de contratos son por tiempo determinado o por necesidades eventuales de las empresas. Generalmente, los trabajadores que se encuentran bajo esta relación laboral pueden ser despedidos sin que el empleador incurra en grandes pagos, de modo que los trabajadores terminen usualmente sin derecho a ninguna indemnización. Esta última característica podría explicar el dramático crecimiento de esta modalidad laboral. El segundo grupo considera a los contratos de temporada, es decir, aquellos que en razón de la costumbre o de la contratación colectiva se celebran entre una empresa o empleador y un trabajador o grupo de trabajadores para que realicen trabajos cíclicos o periódicos, en razón de la naturaleza discontinua de sus labores, gozando estos contratos de estabilidad, entendida como el derecho de los trabajadores a ser llamados a prestar sus servicios en cada temporada que se requieran. Este tipo de contrato, a diferencia de los anteriores, implica la obligación por parte del empleador de renovar el contrato por las siguientes temporadas; caso contrario, el empleador debe incurrir en el pago de indemnización [13].

¹En el año 2007, el INEC introdujo un cambio metodológico en la forma de calcular la tasa de desempleo [17].

Por otro lado, se debe tener presente que en el año 2000 se dictó la Ley para la Transformación Económica del Ecuador “Trole 1” [29], que incorporó la contratación por horas, además de implementar al dólar como moneda de curso legal. Otra de las reformas laborales importantes fue la tercerización, introducida en 1988 con posteriores reformas en el 2006, que no es más que una intermediación de empleo, para formalizar un contrato de trabajo entre empleadores y empleados, para brindar servicios de personal a terceras personas.

Centrándose en el aspecto de la tercerización, no ha sido una reforma positiva para la creación y la calidad del empleo², debido a que los contratos temporales han presentado una evolución progresiva a lo largo de los tres últimos años de análisis. Sin embargo, se presenta un notable decrecimiento a partir de que la Asamblea Constituyente del año 2008, mediante el Mandato Constituyente No. 8 [31]³, en su parte esencial, elimina y prohíbe la tercerización laboral. Así, a partir de este mandato, se produce un decrecimiento en la temporalidad laboral. Pero no se evidencia un efecto progresivo en el largo plazo, pues se presenta una tendencia creciente que se mantiene a lo largo del año 2009. Esto podría deberse a que las empresas ecuatorianas no están en la capacidad de brindar estabilidad a sus trabajadores, ni brindar posibilidades de formación y promoción interna, generándose altas tasas de rotación, menor antigüedad laboral y salarios más bajos para este grupo de asalariados; es decir, existe un desajuste educativo y de movilidad laboral [26].

De acuerdo a la realidad ecuatoriana señalada, el crecimiento dentro del mercado laboral se ha basado en la precarización al trabajador, olvidando que la productividad de los empleados depende no sólo de su aptitud y de la inversión que se realiza en ellos, tanto dentro como fuera del puesto de trabajo, sino también de su motivación y de la intensidad de su esfuerzo [8]. La teoría del capital humano señala que la inversión en la formación de los trabajadores hace que este se convierta en un factor para disminuir el desempleo; es por esto que se pretende verificar si, la formación es un factor influyente al momento de tener un contrato indefinido en el Ecuador. Además, es preciso recalcar la hipótesis de que los trabajadores que se encuentran en período de prueba son más eficientes,

²De acuerdo al Ec. Rafael Correa, presidente de la República del Ecuador [5].

³Eliminación y prohibición de la tercerización, intermediación laboral, contratación laboral por horas y cualquier forma de precarización de las relaciones de trabajo.

de que existe una auto preparación, de que los empleadores tienen un conocimiento previo de sus trabajadores al conocer las capacidades y habilidades del trabajador frente a las funciones que debe desempeñar [30].

Por lo expuesto, este trabajo tiene dos objetivos. El primero consiste en identificar las características primordiales que afectan o determinan que una persona tenga un contrato temporal y, de esta manera, poder influenciar en dichas características. El segundo, consiste en determinar en qué medida la temporalidad es un fenómeno que depende tanto de las características de los trabajadores, como de la rama de actividad de las empresas.

El artículo se estructura como sigue. En la sección 2, se expone la metodología y las especificaciones econométricas que servirán de base para las estimaciones que se realizarán. En la sección 3, se presenta el modelo desarrollado que permite el contraste entre la técnica de estimación empleada, los datos y variables utilizadas. En la sección 4, se señalan las conclusiones más relevantes. Al final, se presenta un anexo en el que se detalla el análisis descriptivo de los datos utilizados.

3.2. Metodología

En esta sección se plantean las especificaciones econométricas necesarias para analizar los factores determinantes al momento de tener un contrato temporal, en contraposición con el contrato indefinido.

La modelización estadística con datos binarios tiene como fin adaptar las herramientas de regresión convencionales para relacionar la variable respuesta (Y) con variables explicativas (X_1, X_2, \dots, X_n) [20].

Dentro de este contexto, los modelos más habituales son: el modelo Logit y el modelo Probit, con la diferencia que el uno utiliza la función de distribución logística y el otro la normal tipificada [6]. En este estudio, se utilizará el modelo Logit por ser el modelo que más se ha potenciado en la aplicación de los estudios empíricos [32].

Para cuantificar la influencia que ejercen las variables explicativas, en la probabilidad de tener una relación laboral temporal frente a una indefinida, se es-

pecifica el siguiente modelo Logit:

$$\Pr(Y = y_j) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m, \quad (3.1)$$

donde

$$Y = \begin{cases} 1 & \text{contrato temporal,} \\ 0 & \text{contrato indefinido} \end{cases}$$

y

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}. \quad (3.2)$$

La ecuación (3.1) representa una función de distribución de probabilidad; en consecuencia, toma sus valores entre 0 y 1. Para estimar los parámetros del modelo, se utiliza el método de máxima verosimilitud.

Tras estimar el modelo, en el siguiente paso se comprobará la significación estadística de cada uno de los coeficientes de la regresión; para ello, se emplea básicamente el estadístico de Wald⁴, ya que nos permite evaluar de forma individual la significancia de la covariable introducida.

Para la interpretación de los parámetros β sin conocer la magnitud de los mismos, ya que no es un modelo lineal, el signo indican que, si el estimador es positivo, incrementos en la variable asociada causarían incrementos en la probabilidad de tener un contrato temporal ($\Pr(Y=1)$). Y si el estimador es negativo, significará que incrementos en la variable asociada causarían disminuciones en la probabilidad de tener un contrato temporal ($\Pr(Y=1)$).

3.3. Modelo

En esta sección se presenta un modelo que permite contrastar la técnica de estimación empleada con los datos y las variables utilizadas. Para lograr los objetivos de esta investigación, que son:

- i Determinar las características que afectan el tener un contrato temporal y,
- ii Determinar en qué medida la temporalidad depende de la rama de actividad.

⁴Estadístico de Wald: Contrasta la hipótesis de que un coeficiente aislado es distinto de 0 y, sigue una distribución normal de media 0 y varianza 1.

3.3.1. Datos

La información estadística utilizada en el análisis proviene de la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) desde el II trimestre del año 2007 al II trimestre del año 2010 del área urbana [24], con las siguientes restricciones en la base:

1. Se considera a las personas de 15 a 65 años de edad, debido a que los Artículos 82 y 86 del Código de la Niñez y Adolescencia [27] señalan que la edad mínima para realizar un contrato laboral se fija en 15 años para todo tipo de trabajo; y, por otro lado, se establece hasta los 65 años de edad, pues es la edad mínima en la que los empleados pueden jubilarse.
2. Se seleccionaron a los individuos que están trabajando y tienen un contrato laboral. Además, se considera a individuos únicos en los diferentes períodos de tiempo; es decir, si un individuo Z fue encuestado en el segundo trimestre del año 2007 y el mismo individuo Z es encuestado nuevamente en el segundo trimestre del año 2008, se tomará la información más actual, es decir, la del segundo trimestre del año 2008.
3. Inicialmente, se contó con una muestra conformada por 586 018 observaciones a nivel nacional. Una vez eliminados los valores de las variables que no se encuentran en el dominio del estudio especificado en los numerales anteriores, la muestra final con la que se realizó el estudio es de 66 655 trabajadores del área urbana con contrato laboral.

Dentro de esta muestra, la composición o estructura de las variables de control que se incorporarán en el modelo toman en cuenta características personales, laborales y geográficas, ya que se pretende verificar si estas son determinantes al momento de tener un contrato temporal. La selección de las variables independientes se basa, en primera instancia, en estudios precedentes que se han realizado en España [10], [11] y [18], pues, a nivel de Ecuador, no se ha encontrado un estudio previo sobre el tema referente.

En segunda instancia, el fundamento para determinar qué variables independientes son las que mejor explican la variable dependiente de dos valores o bivalente (1 tener un contrato temporal, 0 tener un contrato fijo), se llevó a cabo

una serie de submodelos, evaluando la confusión e iteración [4], evitando de esta manera tener resultados estadísticos sesgados; de esta manera cada variable que integra el modelo multivariado es relevante en el estudio analizado, ya que tiene un aporte significativo.

Características Personales

Dentro de las características personales que se analizan, tenemos: el *Sexo*, la *Edad*, el *Nivel de Instrucción* y el *Estado Civil*, con el objeto de analizar diferencias existentes entre las mismas.

1. El porcentaje de individuos que son hombres es el 60,9% y las mujeres el 39,1%; es así que la mayoría de los datos a ser analizados corresponden a individuos con un contrato laboral entre 15 y 65 años de edad que son hombres, reflejándose una diferencia de género en la determinación de contratos. Estos datos se muestran en la tabla 3.1.

Sexo	Número	Porcentaje
Hombre	40 599	60,90
Mujer	26 056	39,09
Total	66 655	100

Tabla 3.1: Frecuencias y Porcentajes para la variable *Sexo*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

2. De los individuos con un contrato laboral se puede observar, en la tabla 3.2, que un 76,74% se encuentra entre personas de 15 a 45 años de edad y que tan solo un 7,65% corresponden a personas entre 56 y 65 años de edad, evidenciándose que, en el Ecuador, la población económicamente activa se encuentra concentrada en personas jóvenes.

Rango de Edad	Número	Porcentaje
15-25	18 586	27,88
26-35	18 357	27,54
36-45	14 213	21,32
46-55	10 400	15,60
56-65	5 099	7,65
Total	66 655	100

Tabla 3.2: Frecuencias y Porcentajes para la variable *Rango de Edad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

3. En la tabla 3.3, se puede observar la estructura de la muestra con respecto al *Nivel de Instrucción* que se maneja de manera original, concentrándose los porcentajes más altos en personas con educación secundaria y personas con educación superior universitaria en un 33,54% y un 30,38%, respectivamente, indicando que, en el Ecuador, existen dos grupos de concentración en los que existe, por un lado, personas que únicamente tienen acceso a una educación básica, por el otro, personas que tienen una educación superior y a la vez el nivel de individuos que tienen post-grado es mínimo.

Nivel de instrucción	Número	Porcentaje
Ninguno	819	1,23
Centro de alfabetización	171	0,26
Primaria	14 695	22,05
Educación Básica	2 962	4,44
Secundaria	22 359	33,54
Educación Media	3 047	4,57
Superior no universitaria	974	1,46
Superior Universitaria	20 248	30,38
Post-grado	1 380	2,07
Total	66 655	100

Tabla 3.3: Frecuencias y Porcentajes para la variable *Nivel de Instrucción*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

4. La clasificación que se presenta en la variable *Estado Civil* es amplia, determinándose seis estados civiles. Las categorías de Separado, Divorciado y Viudo representan, en conjunto, un 11,36%, siendo así más bajo que las otras categorías, como se puede observar en la tabla 3.4. Se evidencia que

en el Ecuador la cantidad de personas casadas es mayor que los otros estados civiles.

Estado Civil	Número	Porcentaje
Casado(a)	24 035	36,06
Soltero(a)	23 605	35,41
Unión libre	11 440	17,16
Separado(a)	4 800	7,20
Divorciado(a)	1 768	2,65
Viudo(a)	1 007	1,51
Total	66 655	100

Tabla 3.4: Frecuencias y Porcentajes para la variable *Estado Civil*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Características Laborales

Las características laborales que se analizan son: la *Ocupación*, *Antigüedad*, *Jornada Laboral*, *Tamaño de Establecimiento* y *Capacitación*, con el objeto de analizar si el tipo de empleo en el cual se encuentra la persona es un factor que tiene influencia en el tipo de contrato.

1. En la muestra a ser analizada, se observa que la mayor parte de los trabajadores se encuentran en el sector privado (60,55%), mientras que se refleja un nivel menor al 1% para empleados tercerizados, como se observa en la tabla 3.5. Esto nos indica que el sector privado es el que ofrece mayor empleo. Cabe recalcar que la ocupación de Jornalero o peón hace referencia a personas que se dedican a tareas sencillas y rudimentarias; se los identifica como aquellas personas que pactan su remuneración por día o jornada independientemente de la periodicidad de pago [25]. La clasificación que se toma en cuenta dentro de la variable *Ocupación* es la definida por el INEC en la ENEMDU, que permite abarcar la realidad del mercado de trabajo dentro de las ocupaciones que se desempeñan en el Ecuador. Se observa en la tabla A.5 (ver página 81), que existen algunos contratos a jornal (2%) que no son contratos temporales; hecho que puede ser explicado, por las características culturales de la región Costa, pues es muy común que trabajadores de las haciendas bananeras laboren de manera continua sin

tener establecido de forma legal un contrato indefinido. Sin embargo, por el tiempo que llevan realizando esta actividad, en muchos casos años, en la encuesta señalan tener un contrato indefinido.

Ocupación	Número	Porcentaje
Empleado privado	40 359	60,55
Empleado de gobierno	13 487	20,23
Jornalero o peón	7 933	11,90
Empleado Doméstico	4 413	6,62
Empleado terciarizado	463	0,69
Total	66 655	100

Tabla 3.5: Frecuencias y Porcentajes para la variable *Ocupación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

- De los individuos con un contrato laboral, cuya *Antigüedad* es mayor a un año, se tienen que estos representan el 69,27%. Evidenciando que la mayoría de población en el Ecuador se mantiene trabajando al menos más de un año, como se puede observar en la tabla 3.6.

Antigüedad	Número	Porcentaje
Mayor a un año	46 169	69,27
Menor o igual a un año	20 486	30,73
Total	66 655	100

Tabla 3.6: Frecuencias y Porcentajes para la variable *Antigüedad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

- En la tabla 3.7 se puede observar que el 83,87% de los individuos con un contrato laboral tienen una *Jornada Laboral* completa, reflejándose, en este caso, la eliminación de contratos temporales vía mandato [13], por lo cual las personas presenten una mayor propensión en contratos con jornada completa.

Jornada Laboral	Número	Porcentaje
Jornada Completa	55 901	83,87
Jornada Parcial	10 754	16,13
Total	66 655	100

Tabla 3.7: Frecuencias y Porcentajes para la variable *Jornada Laboral*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

4. Es interesante observar en la tabla 3.8, que el 64,59 % de los establecimientos tienen menos de 100 trabajadores, reflejándose así que la mayoría de los centros de trabajo que existen en el Ecuador son pequeños, medianos, o que se encuentran en crecimiento.

Tamaño del establecimiento	Número	Porcentaje
Menos de 100	43 052	64,59
100 y más	23 603	35,41
Total	66 655	100

Tabla 3.8: Frecuencias y Porcentajes para la variable *Tamaño del establecimiento*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

5. En la tabla 3.9 se puede observar que tan solo el 23,57 % de las personas con un contrato reciben cursos de capacitación en el lugar de trabajo, evidenciando que los centros de trabajo ecuatorianos que están en la capacidad económica o física de brindar capacitación a su empleado son limitados, ya que la mayoría de los centros de trabajo del país son pequeños, medianos o se encuentran en expansión.

Capacitación	Número	Porcentaje
Si	15 717	23,57
No	50 938	76,42
Total	66 655	100

Tabla 3.9: Frecuencias y Porcentajes para la variable *Capacitación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Características Geográficas

En la tabla 3.10 se puede observar que, entre Sierra y Costa, se encuentra el 94,07 % de la muestra, en la Amazonía el 5,43 % y en las Zonas no delimitadas, tan sólo se encuentra el 0,50 %. Esta variable es importante para reflejar si existe o no diferencia en relación al lugar de residencia, en la posibilidad de contar con un contrato laboral estable. Regiones geográficas o ciudades autorepresentadas como son: Quito, Cuenca, Ambato pertenecen a la Región Sierra y Machala, Guayaquil pertenecen a la Región Costa.

Ciudades Autorepresentadas y Regiones	Número	Porcentaje
Resto de la Sierra	17 072	25,61
Resto de la Costa	15 547	23,32
Guayaquil	8 830	13,25
Quito	6 857	10,29
Cuenca	5 057	7,59
Machala	4 872	7,31
Ambato	4 465	6,70
Amazonía	3 619	5,43
Zonas no delimitadas*	336	0,50
Total	66 655	100

Tabla 3.10: Frecuencias y Porcentajes para la variable *Ciudades Autorepresentadas y Regiones*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

* Las Zonas no delimitadas se refiere a las zonas de las Golondrinas, Manga del Cura y el Piedrero. Se definen como no delimitadas ya que no pertenecen a ninguna región o provincia de manera específica.

Observaciones para el análisis

- El conocer cómo se encuentran distribuidas las características personales, laborales y geográficas de los individuos a ser analizados, en que tengan un contrato temporal o un contrato indefinido, nos permite conocer las variables con sus respectivas categorías disponibles y, principalmente, determinar las que serán de interés en el análisis.
- Para las variables *Nivel de Instrucción* y *Estado Civil* se realizará una agrupación de las categorías que se presentaron en la sección anterior, para facilitar en la interpretación de resultados. Para tener un nivel de información homogéneo la variable *Nivel de Instrucción* se unifica la información de los individuos que se registra con el sistema tradicional al sistema actual de educación [25]. En la variable *Estado Civil* se realizó una unificación entre Casado y Unión Libre, Divorciado y Separado, pues son categorías que, en principio por su condición se las puede agrupar, dejando a las categorías de Soltero y Viudo por separado, para tener un nivel de información más agregado.

- A las variables *Edad* y *Número de Trabajadores* que, en principio, se presentaban como variables discretas, con el fin de tener un nivel de información que tenga un mayor aporte al análisis desarrollado, se las consideró categóricas. A la variable *Edad* se la dividió en rangos, es así que se empezó probando con una amplitud de 5, resultando no significativa en el modelo, por lo que se realizaron pruebas hasta obtener una amplitud que se ajuste al análisis, siendo igual a 10. En lo que respecta a la variable *Número de Trabajadores*, se la representa en forma dicotómica, de tal manera que cuando el número de trabajadores es menor a 100 se le denomina con 1 y a los de 100 y más trabajadores se le denomina con 0. En el Anexo A se puede observar cómo queda la distribución final de los datos.

3.3.2. Resultados del Modelo

Un análisis preliminar indica diferencias marcadas en grupos de individuos, destacándose, principalmente, características como el nivel de instrucción, la edad, el estado civil, la antigüedad, la capacitación que reciben los empleados; en relación de tener un contrato temporal frente a un contrato indefinido. No cabe duda, que mediante el análisis descriptivo, detallado en el Anexo A, se evidencia que un bajo nivel de instrucción, estar entre los 15 y 35 años de edad, ser soltero, tener una antigüedad menor a un año y no recibir ningún tipo de capacitación por parte del empleador, es el perfil característico de las personas que mantienen un contrato laboral temporal. La evidencia empírica mostrada hace evidente la posibilidad de formular un modelo econométrico que permita conocer cuáles son las características de los asalariados que influyen en tener un contrato temporal o un contrato indefinido.

Los resultados obtenidos a través del modelo econométrico planteado, para averiguar el grado de influencia de cada una de las variables incluidas en el mismo, se muestran en la tabla 3.11. En la citada tabla se explican los coeficientes estimados (β), su error estándar (E.T.), su significación estadística (Sig.) con la prueba de Wald, el Exp (β) y sus intervalos de confianza.

Sobre las características principales, evidenciadas en el análisis descriptivo sobre los individuos con contratos temporales y contratos indefinidos, se seleccionó un individuo de referencia, el cual posee características de protección (por

ejemplo, Nivel de Instrucción Superior Universitaria) para realizar el análisis del modelo; el objeto de esta elección es la de poder tener comparabilidad en el análisis.

Variables en la ecuación ^a								
	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para Exp(B)	
							Inferior	Superior
Constante	-3,440*	0,053	4201,741	1	0,000	0,032		
Sexo								
Hombre	0,175*	0,022	63,286	1	0,000	1,192	1,141	1,244
Edad			469,147	4	0,000			
15-25	0,271*	0,032	72,571	1	0,000	1,312	1,232	1,396
26-35	0,220*	0,028	61,595	1	0,000	1,246	1,180	1,317
46-55	-0,261*	0,033	61,479	1	0,000	0,771	0,722	0,823
56-65	-0,555*	0,044	155,625	1	0,000	0,574	0,526	0,627
Nivel Instrucción			511,308	6	0,000			
Post-grado	0,047***	0,079	0,359	1	0,549	1,048	0,898	1,224
Superior no Universitaria	-0,116***	0,084	1,886	1	0,170	0,891	0,755	1,051
Educación Media	0,196*	0,026	56,528	1	0,000	1,216	1,156	1,280
Educación Básica	0,575*	0,028	430,459	1	0,000	1,778	1,684	1,877
Centro de Alfabetización	0,854*	0,211	16,323	1	0,000	2,349	1,552	3,556
Ninguna	1,035*	0,107	93,065	1	0,000	2,816	2,282	3,475
Estado Civil			65,787	3	0,000			
Viudo	0,223*	0,081	7,604	1	0,006	1,250	1,067	1,464
Soltero	0,129*	0,024	28,153	1	0,000	1,137	1,084	1,193
Separado/Divorciado	0,234*	0,034	47,148	1	0,000	1,264	1,182	1,351
Antigüedad								
Menor igual a un año	1,501*	0,023	4178,587	1	0,000	4,485	4,285	4,693
Ocupación			1923,849	4	0,000			
Empleado Doméstico	0,092**	0,054	2,875	1	0,090	1,097	0,986	1,220
Empleado/Obrero Privado	0,098*	0,036	7,623	1	0,006	1,103	1,029	1,183
Empleado Tercerizado	1,692*	0,124	187,192	1	0,000	5,429	4,260	6,918
Jornalero Peón	3,693*	0,093	1588,866	1	0,000	40,150	33,483	48,143
Jornada Laboral								
Jornada Parcial	0,767*	0,029	711,831	1	0,000	2,153	2,035	2,278
Tamaño del Establecimiento								
Menos de 100	0,614*	0,029	457,254	1	0,000	1,847	1,746	1,954
Capacitación								
No recibe	0,884*	0,026	1112,924	1	0,000	2,420	2,298	2,549
Ciudades Autorepresentadas y Regiones			2235,333	8	0,000			
Zonas no delimitadas	1,762*	0,163	117,162	1	0,000	5,824	4,233	8,012
Machala	1,743*	0,052	1123,408	1	0,000	5,714	5,160	6,327
Guayaquil	1,529*	0,045	1149,38	1	0,000	4,614	4,224	5,040
Quito	0,663*	0,047	196,496	1	0,000	1,940	1,768	2,128
Ambato	0,604*	0,052	138,850	1	0,000	1,830	1,654	2,025
Resto de la Sierra	1,054*	0,042	638,166	1	0,000	2,869	2,644	3,114
Resto de la Costa	1,521*	0,043	1276,544	1	0,000	4,578	4,211	4,976
Amazonía	1,216*	0,056	467,668	1	0,000	3,372	3,020	3,765

Tabla 3.11: Resultados de las variables incluidas en el Modelo. Fuente: elaboración propia a partir de la encuesta ENEMDU. ^a El individuo de referencia es mujer, en una edad entre 36 y 45 años, con un nivel de estudios superior universitaria, casado/unión libre, con una antigüedad mayor a un año, empleada de gobierno, con jornada laboral completa, en una empresa de más de 100 trabajadores, recibe cursos de capacitación y reside en Cuenca. (*) Coeficiente significativo al 5 %, (**) al 10 %, (***) no significativos [11].

Análisis

En primer lugar, se puede observar que los coeficientes son significativos con un nivel de confianza del 95% para todas las variables incluidas en el modelo. Cabe recalcar que para determinadas categorías de las siguientes variables: *Nivel de Instrucción* y *Ocupación*, en las categorías Post-grado, Superior no Universitaria y Empleado Doméstico, respectivamente, no son estadísticamente significativas al 95%, ya que el valor de Sig. es mayor a 0,05. Sin embargo, no se eliminan las categorías que resultaron no significativas de estas variables, ya que, en forma global las variables son significativas; es decir, en el caso de las variables categóricas, si alguna resulta significativa, la variable queda incluida en el modelo final, aun cuando el resto no lo sea [16]. Además, debido a que las mismas variables de forma univariante son significativas, se decide dejarlas en el modelo multivariado, ya que aportan información extra dentro del objetivo de la investigación [10].

En segundo lugar, para comprobar si el modelo es globalmente significativo, se evalúa con la prueba del logaritmo del cociente de verosimilitudes [2], [3], en la que se obtiene una significación estadística de (0,00); indicando así que el modelo con las variables incluidas mejora el ajuste de forma significativa.

En tercer lugar, se evalúa el ajuste del modelo mediante una “valoración de prueba diagnosticada”. Esto quiere decir que se comprueba, cómo clasifica el modelo obtenido a nuevos individuos de la muestra, en comparación con la realidad observada. Con el objetivo de evidenciar la clasificación que realiza el modelo con las diez variables predictoras, se presenta la tabla 3.12 de clasificación, en la que se puede apreciar cómo el modelo obtenido clasifica correctamente a 27 934 (de los 35 545) trabajadores temporales ($Y=1$), por lo que su sensibilidad es de 78,60%. Por otra parte, el modelo clasifica correctamente a 23 403 (de los 31 110) trabajadores con contrato indefinido, por lo que la especificidad del modelo es del 75,20%. De forma global se puede decir que se ha clasificado correctamente al 77,0% de los individuos, siendo considerada una predicción aceptable [3]. Además, mediante la curva característica operativa del receptor (curva de COR), la cual discrimina la puntuación óptima de corte para una prueba de detección selectiva, la probabilidad de clasificar correctamente un par de individuos con contrato temporal y contrato indefinido es de 0,85; es decir, el

poder de discriminación del modelo construido es de aproximadamente el 85 %.

Tabla de clasificación					
Observado		Pronosticado			Porcentaje correcto
		Y			
		Contrato Indefinido	Contrato Temporal		
Y	Contrato Indefinido	23403	7707	75,2	
	Contrato Temporal	7611	27934	78,6	
Porcentaje global				77,0	

Tabla 3.12: Tabla de Clasificación. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Finalmente, se presenta el resumen de la validación del modelo.

R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
0,351	0,469

Tabla 3.13: Resumen de la Validación del Modelo. Fuente: elaboración propia a partir de la encuesta ENEMDU.

El **R cuadrado de Cox y Snell** y el **R cuadrado de Nagelkerke** presentado en la tabla 3.13 tiene un valor de 0,351 y 0,469, respectivamente, indicando que el 35,10% y el 46,9% de la variación de la variable dependiente es explicada por las variables incluidas en el modelo. Cabe indicar que el R cuadrado de Nagelkerke es una versión corregida del R cuadrado de Cox y Snell. Además, según [3, 18] se puede considerar que los valores obtenidos representan un buen ajuste en la estimación del modelo.

Interpretación de Resultados

La interpretación de resultados obtenidos, mostrados en la tabla 3.11, se basarán en las características personales, laborales y geográficas consideradas en este estudio.

Las características personales que afectan positivamente la probabilidad de aumentar la obtención de una relación laboral temporal son: que un individuo tenga nivel de educación media, o educación básica, o centro de alfabetización, o ninguna; que sea hombre; que tenga una edad entre 15 y 35 años. Por otro lado, las características que posibilitan una disminución en la probabilidad de

tener un contrato temporal son: que tenga un nivel de instrucción superior, sea mujer, casada y con una edad comprendida entre los 46 y 65 años de edad.

Las características laborales que conducen a aumentar asalariados temporales son: que tenga una jornada laboral parcial, que tenga una experiencia menor o igual a un año de trabajo, que este trabajando en una empresa de menos de 100 trabajadores, que sea jornalero o peón y que no reciba cursos de capacitación. Por el contrario, las características que influyen a disminuir la probabilidad del contrato temporal son: que tenga una jornada laboral completa, que tenga una experiencia mayor que un año, que se encuentre en una empresa de más de 100 trabajadores, que reciba cursos de capacitación y que desarrolle su trabajo dentro del sector público.

De acuerdo a las ciudades autorepresentadas y regiones que se incorporaron en el modelo, se puede notar que en: Guayaquil, Machala, resto de la Costa, Amazonía y Zonas no delimitadas, el contrato temporal tienen mayor probabilidad de ocurrir. Por el contrario, en Quito, Ambato, Cuenca y resto de la Sierra esta probabilidad disminuye.

Escenarios del Modelo

En esta sección, a partir de la base de estudio, se analizan tres diferentes escenarios; considerando los factores de protección y de riesgo, que presentan tres individuos seleccionados. Cabe indicar que el valor $Exp(\beta)$, obtenido para cada una de las variables utilizadas en el modelo, será el indicador de si ésta, es o no una variable de riesgo.

El estudio y planteamiento de estos escenarios, tiene por objeto observar la bondad de las estimaciones realizadas. A continuación, se presenta en la tabla 3.14 los resultados para cada escenario. Los escenarios propuestos son: Pesimista, Moderado y Optimista.

- **Escenario Pesimista:** si se consideran las características menos favorables que puede tener un asalariado, las mismas serían las siguientes: un individuo hombre, entre 15 y 25 años, ningún nivel de instrucción, separado o divorciado, con una antigüedad menor igual que un año, jornalero o peón, con jornada parcial, en un establecimiento de menos de 100 trabaja-

dores, no recibe cursos de capacitación y reside en las zonas no delimitadas del Ecuador. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,998 de tener un contrato temporal, (véase tabla 3.14).

- **Escenario Moderado:** si se considera condiciones aceptables de un asalariado en el mercado laboral, las mismas serían las siguientes: un individuo hombre entre 26 y 35 años de edad, con un nivel de instrucción superior universitaria, soltero, empleado privado, con una antigüedad mayor que un año, con jornada completa, en un establecimiento de más de 100 trabajadores, que recibe capacitación y que habita en la ciudad de Guayaquil. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,216 de tener un contrato temporal, (véase tabla 3.14).
- **Escenario Optimista:** si se considera condiciones relativamente buenas y estables dentro del mercado laboral, el escenario optimista debería ser el siguiente: un individuo mujer, entre 56 y 65 años de edad, con postgrado, casada, empleada pública, con una antigüedad mayor a uno año, con jornada completa, en un establecimiento de más de 100 trabajadores, si recibe cursos de capacitación y habita en la ciudad de Cuenca. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,018 de tener un contrato temporal, (véase tabla 3.14).

Características del Individuo		Escenarios		
		Pesimista	Moderado	Optimista
Sexo	Hombre	1	1	0
	Mujer	0	0	1
Edad	15-25	1	0	0
	26-35	0	1	0
	36-45	0	0	0
	46-55	0	0	0
	56-65	0	0	1
Nivel de Instrucción	Post-grado	0	0	1
	Superior no universitaria	0	0	0
	Superior Universitaria	0	1	0
	Educación Media	0	0	0
	Educación Básica	0	0	0
	Centro de Alfabetización	0	0	0
Estado Civil	Ninguna	1	0	0
	Viudo	0	1	0
	Soltero	0	1	0
	Separado/Divorciado	1	0	0
Antigüedad	Casado/Unión Libre	0	0	1
	Mayor a un año	0	1	1
Ocupación	Menor igual a 1 año	1	0	0
	Empleado Doméstico	0	0	1
	Empleado Privado	0	1	0
	Empleado Público	0	0	1
	Empleado Tercerizado	0	0	0
Jornada Laboral	Jornalero Peón	1	0	0
	Jornada Completa	0	1	1
Tamaño del Establecimiento	Jornada Parcial	1	0	0
	Menos de 100	1	0	0
Capacitación	Más de 100	0	1	1
	Si recibe	0	1	1
Ciudades Autorepresentadas y Regiones	No recibe	1	0	0
	Zonas no delimitadas	1	0	0
	Machala	0	0	0
	Guayaquil	0	1	0
	Quito	0	0	0
	Ambato	0	0	0
	Cuenca	0	0	1
	Resto de la Sierra	0	0	0
	Resto de la Costa	0	0	0
Amazonía	0	0	0	
$Pr(Y = 1)$		0,998	0,216	0,018

Tabla 3.14: Escenarios para el Modelo. Fuente: elaboración propia a partir de la encuesta ENEM-DU.

En los escenarios desarrollados, la distinción del sexo marca una diferencia; es decir, un hombre con las características que se señalan en los tres diferentes escenarios presenta una probabilidad de tener un contrato temporal mayor que la de mujeres.

Comprobación de la hipótesis (i),

Se determinan las características que afectan el tener un contrato temporal.

Las características como: el *Nivel de Educación*, el tipo de *Ocupación*, la *Antigüedad* y la *Jornada Laboral* son variables que, al modificar su estado, afectan en mayor medida la probabilidad de obtener un contrato temporal.

En particular, cuando un individuo es jornalero o peón es el factor de riesgo que más influye en la obtención de tener un contrato temporal. Esto se observa en el escenario pesimista, con las condiciones establecidas se obtuvo una probabilidad alta en la predicción del modelo. Las Zonas no delimitadas son las que presentan una mayor probabilidad de temporalidad y Cuenca es la ciudad de menor probabilidad.

3.3.3. Análisis por Rama de Actividad

El objetivo de este análisis es identificar las características relevantes de cada Rama de Actividad; para, de esta manera, conocer cuál es el grupo de individuos predominante en cada rama y así tener un aporte que permita evidenciar y constatar las razones de la existencia del contrato temporal en cada una de ellas.

Para lograr el objetivo señalado, se realizó un análisis desagregado de acuerdo a la Clasificación Internacional Industrial Uniforme (CIIU) a un dígito, en la cual se consideran 17 Ramas de Actividad que permiten clasificar al establecimiento donde se encuentra el empleado. Cabe señalar que los datos existentes para el sector "Q. Organizaciones y órganos extraterritoriales" no son suficientes para realizar el análisis, pues la única información que se posee es la de 7 individuos a lo largo del período de análisis, y su representación porcentual dentro del conjunto de ramas de actividad consideradas es apenas el 0,01 %, (véase la figura 3.1).

En la figura 3.1, se puede observar la distribución de las distintas Ramas de Actividad, dentro del universo considerado para el estudio.

A continuación, se presentan los resultados obtenidos, enfocándose, principalmente, en los *Rangos de Edad*, en el *Nivel de Instrucción*, en el tipo de *Ocupación*, y en las *Ciudades Autorepresentadas y Regiones*. Sin embargo, no se prestará mayor atención a las variables, *Sexo*, *Antigüedad*, *Jornada Laboral*, *Tamaño del Establecimiento* y *Capacitación*, pues son variables que, al aplicarlas en el modelo por Rama de Actividad, en general, no presentan variaciones sobre cada una de ellas, es decir, estas variables tienen el mismo efecto que se tiene en el modelo presentado en la tabla 3.11.

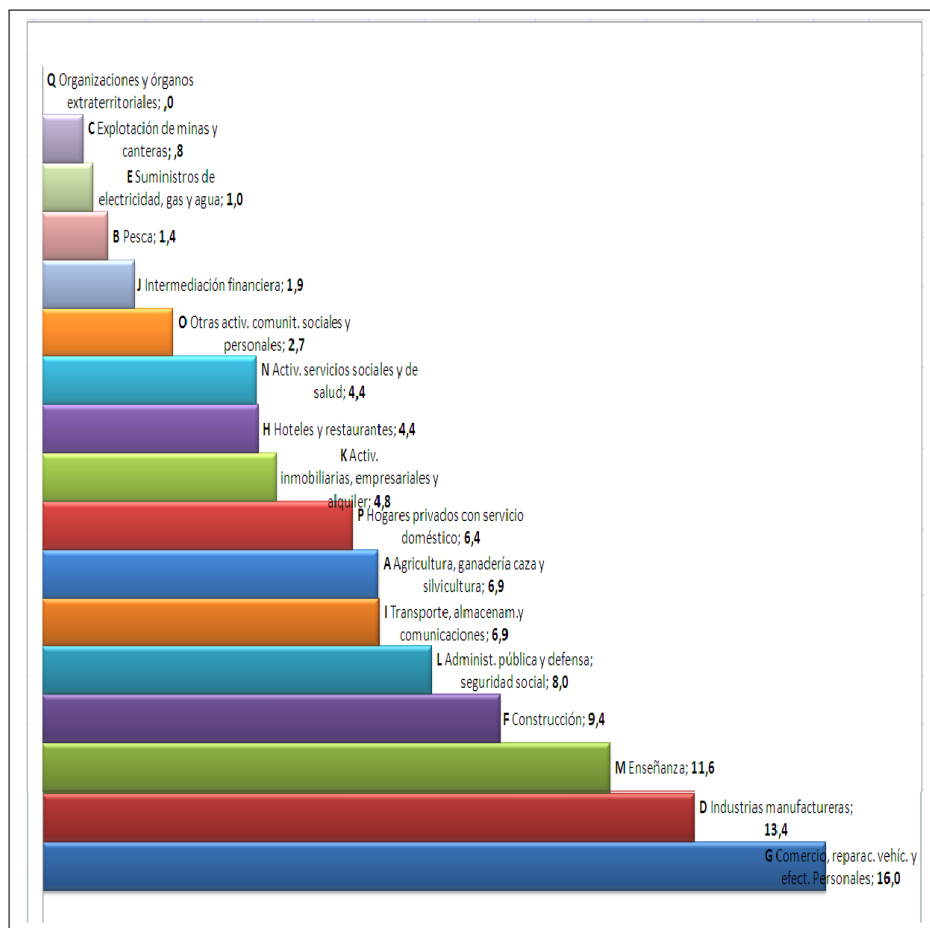


Figura 3.1: Distribución porcentual de individuos por Rama de Actividad. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Es necesario tomar en cuenta que los resultados que se presentan son realizados en base a una muestra, por lo que la agrupación de los individuos en determinadas ciudades representarían el comportamiento general de la región.

■ A. Agricultura, ganadería, caza y silvicultura

Rama de Actividad	Características		Influencia
A.Agricultura, ganadería, caza y silvicultura	Rango de Edad	15-25	Positiva
		26-35	Positiva
	Nivel de Instrucción	Educación Media	Positiva
		Educación Básica	Positiva
Ninguna		Positiva	
Ocupación	Empleado/Obrero Privado	Negativa	
Ciudades Autorepresentadas y Regiones	Machala	Positiva	
	Amazonía	Positiva	
	Zonas no delimitadas	Positiva	
	Guayaquil	Positiva	
	Resto de la Sierra	Positiva	
	Resto de la Costa	Positiva	

Tabla 3.15: Caracterización de los individuos con contrato temporal en la Rama A. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.15 se observa que el factor de protección determinante es el tipo de *Ocupación* que desempeñen; es así, individuos con ocupación de empleado privado dentro de esta rama de actividad tienen menos propensión de tener un contrato temporal. Por otro lado, los factores de riesgos son: la *Edad*, el *Nivel de Instrucción* y el lugar donde residan, por ejemplo, un individuo entre 15 y 35 años de edad, con una educación media, o básica, o ningún tipo de educación, y que residan en Machala, o Guayaquil, o Resto de la Costa, o Resto de la Sierra, o Amazonía, o Zonas no delimitadas, es más propenso a tener un contrato temporal.

■ B. Pesca

Rama de Actividad	Características		Influencia
B.Pesca	Rango de Edad	15-25	Positiva
	Nivel de Instrucción	Educación Básica	Positiva
	Ocupación	Empleado/Obrero Privado	Negativa

Tabla 3.16: Caracterización de los individuos con contrato temporal en la Rama B. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.16 se puede notar que los factores de riesgo son: el *Nivel de Instrucción* y la *Edad*; es así que individuos de 15 a 25 años con educación básica, presentan una mayor propensión a tener un contrato temporal.

Por otro lado, el factor de protección determinante es el tipo de *Ocupación*, pues individuos con empleo privado dentro de esta rama de actividad fomentan la disminución de la probabilidad de tener un contrato temporal. En esta rama de actividad no se puede determinar en qué región o ciudad existe una mayor probabilidad de una relación laboral temporal, debido a que este sector económico no se desarrolla en las ciudades que abarca la representación del universo considerado.

■ C. Explotación de minas y canteras

Rama de Actividad	Características		Influencia
C. Explotación de minas y canteras	Rango de Edad	15-25	Positiva
		46-55	Negativa
	Nivel de Instrucción	Educación Básica	Positiva
	Ocupación	Empleado/Obrero Privado	Negativa

Tabla 3.17: Caracterización de los individuos con contrato temporal en la Rama C. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.17 se observa que entre los factores de riesgo se tiene la *Edad* y el *Nivel de Instrucción*; es así que un individuo de 15 a 25 años de edad, con un nivel de educación básica, presenta una mayor propensión de tener un contrato temporal. Por otro lado, como factor de protección se encuentra la edad, pues un individuo entre 46 y 55 años de edad tiene menos propensión de tener un contrato temporal. Al igual que en la rama de actividad Pesca, en el universo considerado no es posible establecer en que región o ciudad se aumentaría el contrato temporal.

■ D. Industrias manufactureras

En la tabla 3.18 se puede notar que el factor de protección determinante es la *Edad*, personas entre 46 y 65 años de edad dentro de esta rama de actividad tienen menos propensión de tener un contrato temporal. Por otro lado, los factores de riesgo que se presentan son el bajo nivel de instrucción, ser soltero, separado o divorciado, ser empleado tercerizado o jornalero peón y que resida en Machala, o Guayaquil, o Resto de la Costa, o Quito, o Ambato, o Resto de la Sierra, o Amazonía, siendo Machala y el Resto de la Costa los que presentan una mayor propensión de tener un contrato temporal.

Rama de Actividad	Características		Influencia
D. Industrias manufactureras	Rango Edad	46-55	Negativa
		56-65	Negativa
	Nivel de Instrucción	Ninguna	Positiva
		Centro de Alfabetización	Positiva
		Educación Básica Educación Media	Positiva
Estado Civil	Soltero	Positivo	
	Separado Divorciado	Positivo	
Ocupación	Empleado Tercerizado	Positiva	
	Jornalero Peón	Positiva	
Ciudades Autorepresentadas y Regiones		Machala	Positiva
		Guayaquil	Positiva
		Quito	Positiva
		Ambato	Positiva
		Resto de la Sierra	Positiva
		Resto de la Costa Amazonía	Positiva

Tabla 3.18: Caracterización de los individuos con contrato temporal en la Rama D. Fuente: elaboración propia a partir de la encuesta ENEMDU.

■ E. Suministro de electricidad, gas y agua

Rama de Actividad	Características		Influencia
E. Suministro de electricidad, gas y agua	Rango de Edad	15-25	Positiva
		26-35	Positiva
	Nivel de Instrucción	Educación Media	Positiva
	Ocupación	Empleado Tercerizado	Positiva

Tabla 3.19: Caracterización de los individuos con contrato temporal en la Rama E. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.19 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se incrementará para los individuos entre 15 y 35 años de edad, con educación media y que sea un empleado tercerizado. Por otra parte, en esta rama de actividad no se evidencia un factor de protección.

■ F. Construcción

En la tabla 3.20 se puede apreciar que la probabilidad de tener un contrato temporal, en esta rama de actividad, aumenta para los individuos con bajo nivel de instrucción, solteros, con un empleo tercerizado o jornalero y que vivan en Machala, o Guayaquil, o Resto de la Sierra, o Resto de la Costa, o Amazonía. Por otra parte, la característica de los individuos que posibilitan una disminución en la probabilidad de obtener una relación laboral temporal es que sea empleado público.

Rama de Actividad	Características		Influencia
F. Construcción	Nivel de Instrucción	Centro de Alfabetización Educación Básica Educación Media	Positiva Positiva Positiva
	Estado Civil	Soltero	Positiva
	Ocupación	Empleado/Obrero Público Empleado Tercerizado Jornalero o Peón	Negativa Positiva Positiva
	Ciudades Autorepresentadas y Regiones	Machala Guayaquil Resto de la Sierra Resto de la Costa Amazonía	Positiva Positiva Positiva Positiva Positiva

Tabla 3.20: Caracterización de los individuos con contrato temporal en la Rama F. Fuente: elaboración propia a partir de la encuesta ENEMDU.

■ G. Comercio al por mayor y menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos

Rama de Actividad	Características		Influencia
G. Comercio al por mayor y menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos	Rango Edad	15-25	Positiva
		46-55	Negativa
	Nivel de Instrucción	Ninguna	Positiva
		Centro de Alfabetización	Positiva
		Educación Básica Educación Media	Positiva Positiva
Estado Civil	Soltero	Positiva	
	Separado/Divorciado	Positiva	
Ocupación	Empleado Tercerizado	Positiva	
	Jornalero o Peón	Positiva	
Ciudades Autorepresentadas y Regiones	Zonas no delimitadas	Positiva	
	Machala	Positiva	
	Guayaquil	Positiva	
	Quito	Negativa	
	Ambato	Negativa	
	Resto de la Sierra	Negativa	
	Resto de la Costa	Positiva	
Amazonía	Positiva		

Tabla 3.21: Caracterización de los individuos con contrato temporal en la Rama G. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.21 se puede apreciar que entre los factores de protección para tener un contrato temporal tenemos: la *Edad* entre 46 a 55 años de *Edad* y vivir en Quito, o Ambato, o en Resto de la Sierra. Por otra parte, como factores de riesgo para tener un contrato temporal tenemos a individuos entre los 15 y 25 años de edad, con un nivel de educación bajo, que sean solteros, o separados, o divorciados, que se desempeñen como empleado

tercerizado, o jornalero o peón y que vivan en Machala, o Guayaquil, o Resto de la Costa, o Amazonía, o Zonas no delimitadas.

■ H. Hoteles y restaurantes

Rama de Actividad	Características		Influencia
H. Hoteles y restaurantes	Rango Edad	15-25	Positiva
	Nivel de Instrucción	Educación Básica	Positiva
	Ciudades Autorepresentadas y Regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Quito	Negativa
		Ambato	Negativa
Resto de la Sierra	Negativa		
Resto de la Costa	Positiva		
Amazonía	Negativa		

Tabla 3.22: Caracterización de los individuos con contrato temporal en la Rama H. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.22 se puede notar que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es por la agrupación de los individuos entre 15 y 25 años de edad, con un nivel de educación básica, y que vivan en Machala, o Guayaquil, o Resto de la Costa, o Zonas no delimitadas. Por otra parte, las características que posibilitan una disminución en la probabilidad de tener un contrato temporal es que vivan en Quito, o Ambato, o Resto de la Sierra, o Amazonía.

■ I. Transporte, almacenamiento y comunicaciones

Rama de Actividad	Características		Influencia
I. Transporte, almacenamiento y comunicaciones	Rango Edad	15-25	Positiva
		26-35	Positiva
	Nivel de Instrucción	Educación Básica	Positiva
		Educación Media	Positiva
	Ocupación	Empleado Tercerizado Jornalero o Peón	Positiva Positiva
	Ciudades Autorepresentadas y Regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
Guayaquil		Negativa	
Quito		Negativa	
Resto de la Sierra		Negativa	
Resto de la Costa	Positiva		
Amazonía	Negativa		

Tabla 3.23: Caracterización de los individuos con contrato temporal en la Rama I. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.23 se puede apreciar que la probabilidad para que incremente

el contrato temporal, en esta rama de actividad, es en los individuos entre los 15 y 35 años de edad, con un nivel de educación básica o media, desarrollándose en un empleo tercerizado o de jornalero peón y que vivan en Machala, o resto de la Costa, o Zonas no delimitadas. Por otra parte, las características que posibilitan la disminución de esta probabilidad, de tener un contrato temporal, es que residan en Guayaquil, o Quito, o Resto de la Sierra, o Amazonía. f

■ J. Intermediación financiera

Rama de Actividad	Características		Influencia
J. Intermediación financiera	Nivel de Instrucción	Educación Básica	Positiva
		Educación Media	Positiva
	Ocupación	Empleado Privado	Negativa
		Empleado Tercerizado	Positiva
Ciudades Autorepresentadas y Regiones	Machala	Quito	Negativa
		Machala	Positiva

Tabla 3.24: Caracterización de los individuos con contrato temporal en la Rama J. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.24 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se ve afectada positivamente por las personas que tenga un nivel de educación básica o media, que sea empleado de tercerizado y que vivan en Machala. Por otra parte, las características que hacen que disminuya la probabilidad de tener un contrato temporal son: ser empleado privado y que vivan en Quito.

■ K. Actividades inmobiliarias, empresariales y de alquiler

Rama de Actividad	Características		Influencia
K. Actividades inmobiliarias, empresariales y de alquiler	Nivel de Instrucción	Educación Básica	Positiva
		Educación Media	Positiva
	Estado Civil	Viudo	Positiva
		Separado/Divorciado Soltero	Positiva Positiva
	Ocupación	Empleado Tercerizado	Positiva
Ciudades Autorepresentadas y Regiones	Machala Guayaquil Quito Resto de la Sierra Resto de la Costa Amazonía	Machala	Positiva
		Guayaquil	Positiva
		Quito	Negativa
		Resto de la Sierra	Negativa
		Resto de la Costa	Positiva
Amazonía	Positiva		

Tabla 3.25: Caracterización de los individuos con contrato temporal en la Rama K. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.25 se puede apreciar que la probabilidad de incrementar el contrato temporal, en esta rama de actividad, es para los individuos con un nivel de educación básica o media, que sean viudos, o separados, o divorciados, o solteros, que tengan un empleo tercerizado y que vivan en Machala, o Guayaquil, o Resto de la Costa. Por otra parte, las características que hacen que disminuya esta probabilidad es que vivan en Quito, o Resto de la Sierra.

■ **L. Administración pública y defensa; planes de seguridad social de afiliación obligatoria**

Rama de Actividad	Características		Influencia
L. Administración pública y defensa; planes de seguridad social de afiliación obligatoria	Rango de Edad	15-25	Positiva
		26-35	Positiva
		56-65	Negativa
	Nivel de Instrucción	Educación Básica	Positiva
	Estado Civil	Separado/ Divorciado Soltero	Positiva Positiva
Ciudades	Machala	Positiva	
Autorepresentadas y	Ambato	Negativa	
Regiones	Resto de la Sierra	Negativa	

Tabla 3.26: Caracterización de los individuos con contrato temporal en la Rama L. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.26 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se ve afectada positivamente para las personas entre 15 y 35 años de edad, con un nivel de educación básica, que sea separado, o divorciado, o soltero y, que vivan en Machala. Por otra parte, las características que posibilitan la disminución de la probabilidad de una relación laboral temporal son: que se encuentre entre los 56 y 65 años de edad y, que vivan en Ambato, o el resto de la Sierra.

■ **M. Enseñanza**

En la tabla 3.27 se puede notar que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es por la agrupación de los individuos entre 15 y 35 años de edad, con ningún nivel de instrucción o centro de alfabetización y que vivan en Machala, o Guayaquil, o el resto de la Costa. Por otra parte, las características que hacen que la probabilidad

Rama de Actividad	Características		Influencia
M. Enseñanza	Rango Edad	15-25 26-35 46-55 56-65	Positiva Positiva Negativa Negativa
	Nivel de Instrucción	Ninguna Centro de Alfabetización	Positiva Positiva
	Ciudades Autorepresentadas y Regiones	Machala Guayaquil Ambato Resto de la Costa Quito	Positiva Positiva Negativa Positiva Negativa

Tabla 3.27: Caracterización de los individuos con contrato temporal en la Rama M. Fuente: elaboración propia a partir de la encuesta ENEMDU.

de obtener un contrato temporal disminuya, son que los individuos tengan una edad entre 46 y 65 años y que vivan en Ambato o Quito.

■ N. Actividades de servicios sociales y de salud

Rama de Actividad	Características		Influencia
N. Actividades de servicios sociales y de salud	Rango Edad	15-25 26-35 46-55 56-65	Positiva Positiva Negativa Negativa
	Nivel de Instrucción	Superior no universitaria Educación Básica	Negativa Positiva
	Ocupación	Empleado Privado	Negativa
	Ciudades Autorepresentadas y Regiones	Machala	Positiva

Tabla 3.28: Caracterización de los individuos con contrato temporal en la Rama N. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.28 se aprecia que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es para los individuos entre 15 y 35 años de edad, con un nivel de educación básica y que vivan en Machala. Por otra parte, las características que hacen que la probabilidad de obtener un contrato temporal disminuya son: que tengan de 46 a 65 años de edad, con un nivel de instrucción superior no universitario y un empleo privado.

■ O. Otras actividades comunitarias sociales y personales de tipo servicios

Rama de Actividad	Características		Influencia
O. Otras actividades comunitarias sociales y personales de tipo servicios	Rango Edad	56-65	Negativa
	Estado Civil	Soltero	Positiva
		Separado/Divorciado	Positiva
	Ocupación	Empleado Tercerizado	Positiva
	Ciudades Autorepresentadas y Regiones	Quito	Negativa
		Amazonía	Negativa
Resto de la Sierra		Negativa	
Machala		Positiva	
Guayaquil		Positiva	
	Resto de la Costa	Positiva	

Tabla 3.29: Caracterización de los individuos con contrato temporal en la Rama O. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.29 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se incrementa para las personas solteras, o separadas, o divorciadas y, que vivan en Machala, o Guayaquil, o el resto de la Costa. Por otra parte, las características que incentivan la disminución de esta probabilidad son: que tenga una edad entre 56 y 65 años y que vivan en Quito, o en Amazonía, o resto de la Sierra.

■ P. Hogares privados con servicio doméstico

Rama de Actividad	Características		Influencia
P. Hogares privados con servicio doméstico	Rango Edad	15-25	Positiva
		56-65	Negativa
	Estado Civil	Soltero	Negativa
	Ciudades Autorepresentadas y Regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Quito	Positiva
		Ambato	Positiva
		Resto de la Sierra	Positiva
Resto de la Costa		Positiva	
Amazonía	Positiva		

Tabla 3.30: Caracterización de los individuos con contrato temporal en la Rama P. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 3.30 se puede apreciar que la probabilidad de disminuir la contratación temporal, en esta rama de actividad, es para los individuos entre 56 y 65 años de edad, soltero. Por otra parte, las características que permiten aumentar la probabilidad de tener un contrato temporal son: que tenga una edad entre 15 y 25 años, y que vivan en Quito, o Ambato, o Resto de la Sierra, o Machala, o Guayaquil, o Resto de la Costa, o Amazonía, o Zonas no delimitadas. Además, es necesario conocer que en esta rama se agrupan las empleadas(o) domésticas(o), y que por tal razón el establecimiento no sobrepasa en ninguno de los casos a 100 trabajadores.

Comprobación de la hipótesis (ii),

Determinar en qué medida la temporalidad depende de la rama de actividad.

De acuerdo a la Rama de Actividad, los factores de riesgo y protección cambian, como se puede observar en la tabla 3.31, denotando a (+) como el factor de riesgo y (-) como el factor de protección.

Características		Rama de Actividad															
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Edad	15-25	+	+	+		+		+	+	+			+	+	+		+
	26-35	+				+				+			+	+	+		
	36-45																
	46-55			-	-			-						-	-		
	56-65				-								-	-	-	-	-
Nivel de Instrucción	Post-grado																
	Superior no universitaria															-	
	Educación Media	+			+	+	+	+		+	+	+					
	Educación Básica	+	+	+	+		+	+	+	+	+	+	+		+		
	Centro de Alfabetización				+		+	+							+		
Ninguna	+			+			+							+			
Estado Civil	Viudo											+					
	Soltero				+		+	+				+	+			+	-
	Separado/Divorciado				+			+				+	+			+	
	Casado / Unión Libre																
Ocupación	Empleado Doméstico																
	Empleado Privado	-	-	-												-	
	Empleado Público							-									
	Empleado Tercerizado				+	+	+	+		+	+	+				+	
	Jornalero o Peón				+		+	+		+							
Ciudades Autorepresentadas y Regiones	Resto de la Sierra	+			+		+	-	-	-		-	-			-	+
	Resto de la Costa	+			+		+	+	+	+		+		+		+	+
	Guayaquil	+			+		+	+	+	-		+		+		+	+
	Quito				+			-	-	-	-	-		-		-	+
	Cuenca																
	Machala	+			+		+	+	+	+	+	+	+	+	+	+	
	Ambato				+			-	-				-	-			+
	Amazonía	+			+		+	+	-	-		+				-	+
Zonas no delimitadas	+						+	+	+							+	

Tabla 3.31: Factores de riesgo y protección por Rama de Actividad. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Por ejemplo para la la Rama de Actividad B tenemos que los factores de riesgo serían: ser una persona entre 15 y 25 años de edad y tener educación básica; por el contrario, el factor de protección sería: ser empleado privado. El análisis es similar para el resto de Ramas de Actividad

3.4. Conclusiones

A partir de datos de la Encuesta de Empleo, Desempleo y Subempleo del Instituto Nacional de Estadísticas y Censos del Ecuador y de la aplicación de un modelo de regresión logística, usando la desagregación de las características personales, laborales y geográficas de los individuos, este trabajo ayuda a comprender la realidad laboral ecuatoriana, pues permite detectar las diferencias en el acceso al empleo temporal según la rama de actividad económica, usando la desagregación de las características personales, laborales y geográficas de los individuos.

Los resultados obtenidos han mostrado que la temporalidad en Ecuador no sólo depende de la estructura productiva existente, sino que es un fenómeno asociado a las características de los trabajadores y a nuestra cultura empresarial. Resaltan la influencia que tiene el nivel de instrucción, el tiempo de trabajo, el tipo de jornada laboral, la capacitación al momento de obtener un contrato temporal, y el hecho de ser hombre o mujer. Los hombres presentan una mayor propensión de tener un contrato temporal. Esto podría explicarse por el hecho de que existen tareas sencillas y rudimentarias, como las que exigen los cultivos agrícolas, la cría de ganado, la pesca, la caza, la industria y construcción que requieran la utilización de herramientas manuales y, a menudo, un esfuerzo físico considerable.

Se observa que el tipo de ocupación es una variable influyente sobre la determinación del tipo de contrato que la persona tenga o pueda llegar a obtener, evidenciando que si una persona tiene una ocupación de jornalero o peón, es una condición que afecta fuertemente a la probabilidad de tener un contrato temporal. Esta situación implica un estancamiento de la situación laboral de las personas que se desenvuelven dentro de esta ocupación, logrando que no puedan tener un nivel salarial ascendente, afectando con esto el nivel económico de las familias.

En referencia a las ciudades autorepresentadas y regiones, se obtiene que en la Región Costa, particularmente en las ciudades de Machala y Guayaquil se concentran los contratos temporales. Una posible explicación sería el desarrollo en las actividades de: agricultura, ganadería, caza y silvicultura; industrias manufactureras; hoteles y restaurantes; transporte; almacenamiento y comunicaciones. En las actividades mencionadas se pueden presentar contratos eventuales, por obra o por horas. Generalmente son actividades que se desarrollan por temporadas, es decir, por períodos del año en los cuales tienen mayor desarrollo y surge la necesidad de incrementar el contrato temporal.

En definitiva se ha comprobado a través de un modelo de regresión logística el conocimiento empírico del comportamiento social, dentro del mercado laboral ecuatoriano, poniendo en evidencia cómo se encuentra concentrado el contrato temporal.

Capítulo 4

Actualización: Análisis de Resultados

En este capítulo, se presentará un análisis de los resultados obtenidos al comparar el uso de las dos técnicas estadísticas utilizadas en la clasificación de individuos, como son la regresión logística y la máquina de vectores soporte.

4.1. Modelo Logit

En el capítulo anterior, se presentó un modelo de regresión logística que abarca una muestra que toma en cuenta datos desde junio del 2007 a junio del 2010. Con el objeto de verificar la validez y robustez de los resultados obtenidos y al disponer de una gama más amplia de datos para analizar la contratación temporal en el Ecuador, se comprobará si los resultados que se obtuvieron en el primer estudio cambian significativamente. Además, se busca determinar si con una cantidad menor de datos se puede llegar a resultados confiables. Para ello, el procedimiento llevado a cabo se realiza en tres etapas, las cuales son:

1. Se realiza el modelo logit con la base de diciembre de cada año desde el año 2007 hasta el año 2011; es decir, se generará un modelo logit para cada diciembre, obteniendo así 5 análisis.
2. Se agrupa las bases a manera de cascada por años, en otras palabras, se tiene una primera base que considera los datos de las bases trimestrales

del año 2007, una segunda base que engloba los datos de las bases trimestrales de los años 2007 y 2008 y, así, se repite el proceso hasta tener la base que contenga los datos de las bases trimestrales desde el año 2007 hasta el año 2012. De esta manera se obtienen 7 periodos de análisis.

3. Se compara los resultados que se alcanzan al cambiar la cantidad de individuos que son considerados para la construcción del estudio.

Para la comparación de los estudios se considera la evaluación de la tasa de clasificación y la significancia individual de las variables incluidas.

En primer lugar, se evalúa la tasa de clasificación, en la tabla 4.1 se agrupa el número de datos, la tasa de clasificación global, la sensibilidad, la especificidad y el R^2 Nagelkerke con los que se está trabajando cada estudio.

Periodo	Dic.07	Dic.08	Dic.09	Dic.10	Dic.11	Jun07Dic07	Jun07Dic08	Jun07Dic09	Jun07Jun10	Jun07Dic10	Jun07Dic11	Jun07Jun12
No. datos	10 474	10 708	10 544	11 049	9 782	20 041	35 997	53 276	66 655	73 243	92 099	104 604
Porcentaje Global	76,752	78,960	77,788	77,708	77,223	75,969	76,437	76,729	77,0	77,0	77,1	77,0
Sensibilidad	76,979	80,784	78,065	76,054	72,475	72,441	76,652	78,295	78,6	78,5	77,5	76,9
Especificidad	76,524	76,952	77,474	79,482	81,838	79,339	76,207	74,950	75,2	75,3	76,6	77,1
R^2 Nagelkerke	0,483	0,506	0,502	0,498	0,472	0,460	0,457	0,463	0,469	0,466	,468	0,468

Tabla 4.1: Resultados de clasificación de los estudios. Fuente: elaboración propia a partir de la encuesta ENEMDU. El modelo a jun07jun10 se encuentra en color azul dado que son los resultados del trabajo publicado que sirvió de base para esta tesis

Se puede observar en la tabla 4.1 los siguientes aspectos:

- El número de datos en las bases de diciembre en promedio se mantiene en 10 511 trabajadores con contrato de trabajo; mientras que, cuando se comienzan a fusionar las bases de cada año la muestra de individuos aumenta aproximadamente por año un 31 %.
- El porcentaje de clasificación global, de los modelos realizados, se diferencia entre los modelos individuales y los modelos agregados. Para los modelos individuales, el correspondiente a diciembre de 2008 es el que logra una mayor tasa de clasificación (78,96 %); en tanto que, para los agregados, el de jun07-dic11 es el que obtiene mejores resultados (77,1 %). Este resultado, es mejor al encontrado con la base de jun07jun10 (trabajo publicado) y sigue siendo mejor al encontrado con la base jun07jun12 (77,0 %). Esto último, en principio indicaría que el agregar información no necesariamente mejora los resultados, pues dependerá de la calidad de la misma o posiblemente de la presencia de factores externos.

- La sensibilidad, de los estudios realizados, se diferencia entre los estudios individuales y los estudios agregados. El correspondiente a diciembre 2008 es el que logra el mayor valor (80,78%) para los estudios individuales; mientras que, para los agregados, el de jun07-jun10 es el que tiene mejores resultados(78,6%). En cambio, la menor sensibilidad se obtiene en el modelo de jun07- dic07 (72,44%).
- La especificidad, de los estudios realizados, también se diferencia entre los estudios individuales y los estudios agregados. El correspondiente a diciembre 2011 es el que logra el mayor valor (81,838%) y, por su parte, la menor especificidad se encuentra en los estudios agregados en el de jun07- dic09 (74,95%).
- Por último, el R^2 Nagelkerke es superior a 0,4 únicamente en los estudios de diciembre 2008 y 2009; esto indica, que el 40% de la variación de la variable dependiente es explicada por las variables incluidas en el análisis. De acuerdo a [2] y [18], se puede considerar que los valores obtenidos representan un buen ajuste.

En este sentido, en primer lugar, se puede concluir que los análisis de diciembre 2007, 2009, 2010, 2011 y junio 2007-diciembre 2011 tienen una mejor tasa de clasificación global que el modelo de junio 2007- junio 2010. La diferencia recae en que este último tiene una mayor tasa de sensibilidad, lo que permite tener una mejor clasificación para los individuos que poseen un contrato temporal.

En segundo lugar, se evalúa la significancia individual de las variables incluidas en el estudio. En las tablas 4.2 y 4.3 se explican los coeficientes estimados (β) y su significación estadística (Sig.) para los estudios con el modelo logit desarrollados con las diferentes bases. La significancia estadística de las variables a un nivel de confianza del 95 %, difieren de estudio a estudio; se puede ver que en algunas variables importantes su significancia varía de ser significativas o no serlo y viceversa.

Variables en la ecuación ^a	dic-07		dic-08		dic-09		dic-10		dic-11	
	B	Sig.	B	Sign	B	Sign	B	Sign	B	Sign
Sexo										
Hombre	0,139	0,015	0,232	0,000	0,101	0,075***	0,131	0,017	0,114	0,045
Edad										
15-25	0,204	0,012	0,326	0,000	0,327	0,000	0,469	0,000	0,413	0,000
26-35	0,133	0,059***	0,207	0,003	0,230	0,001	0,401	0,000	0,199	0,006***
46-55	-0,514	0,000	-0,646	0,000	-0,565	0,000	-0,562	0,000	-0,805	0,0000
56-65	-0,269	0,001	-0,216	0,008	-0,302	0,000	-0,486	0,000	-0,347	0,0000
Nivel Instrucción										
Post-grado	0,178	0,432***	0,152	0,469***	0,371	0,046	0,327	0,042	0,196	0,303
Superior no Universitaria	-0,104	0,628***	-0,601	0,011	0,018	0,932***	-0,115	0,579***	0,091	0,664***
Educación Media	0,931	0,000	1,396	0,000	0,933	0,001	0,991	0,000	1,359	0,000
Educación Básica	0,365	0,000	0,335	0,000	0,239	0,000	0,104	0,113***	0,110	0,099***
Centro de Alfabetización	0,628	0,000	0,637	0,000	0,756	0,000	0,518	0,000	0,679	0,000
Ninguna	0,239	0,758***	1,209	0,012	1,349	0,007	0,972	0,107***	0,815	0,285***
Estado Civil										
Viudo	0,386	0,051***	0,019	0,931***	0,127	0,529***	0,071	0,701***	-0,115	0,615***
Soltero	0,200	0,002	0,163	0,011	0,151	0,016	0,045	0,459***	0,123	0,051
Separado/Divorciado	0,202	0,024	0,354	0,000	0,220	0,010	0,165	0,044	0,260	0,002
Antigüedad										
Menor igual a un año	1,278	0,000	1,467	0,000	1,636	0,000	1,686	0,000	1,573	0,000
Ocupación										
Empleado Doméstico	0,771	0,000	0,469	0,001	0,264	0,065***	0,029	0,840***	-0,339	0,028
Empleado/Obrero Privado	0,794	0,000	0,344	0,000	0,178	0,048	-0,235	0,006	-0,424	0,000
Empleado Tercerizado	2,011	0,000	2,362	0,000	1,074	0,011	-0,082	0,824***	0,295	0,449***
Jornalero Peón	4,096	0,000	3,567	0,000	3,906	0,000	3,792	0,000	3,263	0,000
Jornada Laboral										
Jornada Parcial	0,869	0,000	0,674	0,000	0,911	0,000	0,821	0,000	1,015	0,000
Tamaño del Estable.										
Menos de 100	-0,481	0,000	-0,740	0,000	-0,637	0,000	-0,602	0,000	-0,755	0,000
Capacitación										
No recibe	0,843	0,000	0,811	0,000	0,842	0,000	0,905	0,000	0,823	0,000
CCiudad. Autrep. y Reg.										
Zonas no delimitadas	0,677	0,122***	0,715	0,109***	2,488	0,000	2,977	0,000	1,492	0,007
Machala	1,262	0,000	2,310	0,000	1,487	0,000	1,242	0,000	1,616	0,000
Guayaquil	1,088	0,000	1,873	0,000	1,409	0,000	1,915	0,000	1,139	0,000
Quito	0,412	0,001	1,049	0,000	0,539	0,000	0,749	0,000	0,107	0,431***
Ambato	0,480	0,001	0,602	0,000	0,292	0,050***	1,192	0,000	0,686	0,000
Resto de la Sierra	0,852	0,000	1,615	0,000	0,964	0,000	1,392	0,000	1,049	0,000
Resto de la Costa	1,205	0,000	2,101	0,000	1,452	0,000	1,909	0,000	1,021	0,000
Amazonia	1,091	0,000	1,845	0,000	1,024	0,000	1,387	0,000	0,886	0,000
Constante	-3,298	0,000	-3,446	0,000	-2,845	0,000	-2,879	0,000	-2,080	0,000

Tabla 4.2: Resultados de los estudios. Fuente: elaboración propia a partir de la encuesta ENEMDU.
^a El individuo de referencia es mujer, en una edad entre 36 y 45 años, con un nivel de estudios superior universitaria, casado/unión libre, con una antigüedad mayor a un año, empleada de gobierno, con jornada laboral completa, en una empresa de más de 100 trabajadores, recibe cursos de capacitación y reside en Cuenca. (***) Coeficiente no significativo al 5 % [11].

Variables en la ecuación ^a	Jun07Dic07		Jun07Dic08		Jun07Dic09		Jun07Jun10		Jun07Dic10		Jun07Dic11		Jun.07Jun12	
	B	Sign	B	Sign	B	Sign	B	Sign	B	Sig.	B	Sig.	B	Sign
Sexo														
Hombre	0,198	0,000	0,231	0,000	0,179	0,000	0,175	0,000	0,161	0,000	0,158	0,000	0,148	0,000
Edad														
15-25	0,184	0,001	0,261	0,000	0,220	0,000	0,271	0,000	0,283	0,000	0,309	0,000	0,318	0,000
26-35	0,172	0,001	0,173	0,000	0,1644	0,000	0,220	0,000	0,231	0,000	0,234	0,000	0,228	0,000
46-55	-0,525	0,000	-0,542	0,000	-0,6018	0,000	-0,555	0,000	-0,559	0,000	-0,586	0,000	-0,605	0,000
56-65	-0,208	0,000	-0,224	0,000	-0,252	0,000	-0,261	0,000	-0,311	0,000	-0,319	0,000	-0,324	0,000
Nivel Instrucción														
Post-grado	0,031	0,846***	0,058	0,612***	0,089	0,314***	0,047	0,549***	0,046	0,535***	0,0382	0,559***	0,008	0,900***
Superior no Universitaria	-0,048	0,759***	-0,202	0,105***	-0,177	0,073***	-0,116	0,169***	-0,134	0,092***	-0,141	0,049***	-0,158	0,019
Educación Media	0,770	0,000	1,095	0,000	1,025	0,000	1,035	0,000	0,957	0,000	0,961	0,000	0,959	0,000
Educación Básica	0,333	0,000	0,258	0,000	0,202	0,000	0,196	0,000	0,176	0,000	0,150	0,000	0,157	0,000
Centro de Alfabetización	0,582	0,000	0,618	0,000	0,599	0,000	0,575	0,000	0,544	0,000	0,545	0,000	0,553	0,000
Ninguna	0,337	0,664***	0,694	0,029	1,013	0,000	0,854	0,000	0,709	0,001	0,750	0,000	0,755	0,000
Estado Civil														
Viudo	0,474	0,001	0,275	0,013	0,226	0,012	0,223	0,006	0,231	0,002	0,217	0,002	0,219	0,001
Soltero	0,215	0,000	0,139	0,000	0,154	0,000	0,129	0,000	0,103	0,000	0,108	0,000	0,115	0,000
Separado/Divorciado	0,303	0,000	0,339	0,000	0,272	0,000	0,234	0,000	0,214	0,000	0,221	0,000	0,248	0,000
Antigüedad														
Menor igual a un año	1,397	0,000	1,399	0	1,507	0	1,501	0	1,517	0	1,549	0	1,550	0
Ocupación														
Empleado Doméstico	0,524	0,000	0,200	0,006	0,131	0,030	0,092	0,089***	0,058	0,264***	-0,081	0,083***	-0,130	0,003
Empleado/Obrero Privado	0,602	0,000	0,326	0,000	0,197	0,000	0,098	0,006	0,01929	0,568***	-0,131	0,000	-0,183	0,000
Empleado Tercerizado	2,123	0,000	2,169	0,000	1,817	0,000	1,692	0,000	1,482	0,000	1,276	0,000	1,254	0,000
Jornalero Peón	3,99	0,000	3,682	0,000	3,755	0,000	3,693	0	3,674	0	3,658	0	3,658	0
Jornada Laboral														
Jornada Parcial	0,816	0,000	0,773	0,000	0,763	0,000	0,767	0,000	0,766	0,000	0,804	0,000	0,812	0,000
Tamaño del Estable.														
Menos de 100	-0,396	0,000	-0,507	0,000	-0,569	0,000	0,614	0,000	-0,593	0,000	-0,647	0,000	-0,644	0,000
Capacitación														
No recibe	0,834	0,000	0,868	0,000	0,872	0,000	0,884	0,000	0,877	0,000	0,883	0	0,887	0
Ciud. Autrep. y Reg.														
Zonas no delimitadas	0,777	0,004	1,340	0,000	-1,752	0,000	1,762	0,000	1,839	0,000	1,752	0,000	1,624	0,000
Machala	0,806	0,000	1,591	0,000	0,000	0,998***	1,743	0,000	1,731	0,000	1,820	0	1,857	0
Guayaquil	0,662	0,000	1,297	0,000	-0,241	0,162***	1,529	0,000	1,601	0,000	1,593	0	1,579	0
Quito	0,000	0,997***	0,534	0,000	-1,074	0,000	0,663	0,000	0,683	0,000	0,626	0,000	0,600	0,000
Ambato	0,044	0,641***	0,335	0,000	-1,226	0,000	0,604	0,000	0,748	0,000	0,786	0,000	0,897	0,000
Resto de la Sierra	0,539	0,000	0,956	0,000	-0,709	0,000	1,054	0,000	1,136	0,000	1,144	0,000	1,128	0,000
Resto de la Costa	0,892	0,000	1,406	0,000	-0,253	0,139***	1,521	0,000	1,608	0	1,489	0	1,489	0
Amazonia	0,751	0,000	1,069	0,000	-0,547	0,002	1,216	0,000	1,249	0,000	1,269	0,000	1,252	0,000
Constante	-2,976	0,000	-3,039	0,000	-1,160	0,000	-3,439	0	-2,793	0	-2,652	0	-2,619	0

Tabla 4.3: Resultados de los estudios. Fuente: elaboración propia a partir de la encuesta ENEMDU.

^a El individuo de referencia es mujer, en una edad entre 36 y 45 años, con un nivel de estudios superior universitaria, casado/unión libre, con una antigüedad mayor a un año, empleada de gobierno, con jornada laboral completa, en una empresa de más de 100 trabajadores, recibe cursos de capacitación y reside en Cuenca. (***) Coeficiente no significativo al 5% [11].

En las tablas 4.2 y 4.3, se puede observar que:

- En el caso de las características personales: el sexo se torna no significativo en el estudio correspondiente a diciembre 2009. El rango de edad de 26 a 35 es no significativo para los estudios de diciembre 2007 y diciembre 2011. La categoría de *Post-grado* en la mayoría de los estudios resulta ser no significativa a excepción de los estudios de diciembre 2009, diciembre 2010 y, diciembre 2011. La categoría de *Ninguna* es no significativa para los estudios de diciembre 2007, diciembre 2010, diciembre 2011 y Junio - Diciembre 2007. La categoría de *viudo* resulta ser no significativa para los estudios de cada año, es decir, desde diciembre del año 2007 a diciem-

bre de 2011. Finalmente, la categoría de *soltero* resulta no ser significativa únicamente para el estudio de diciembre 2010.

- En el caso de las características laborales: se tiene que la categoría de *empleado doméstico* resulta ser no significativa para los estudios de diciembre 2009, diciembre 2010, Junio 2007- Junio 2010, Junio 2007- Diciembre 2010 y Junio 2007-Diciembre 2011. Por su parte, la categoría *empleado tercerizado* se torna no significativa únicamente para los estudios de diciembre 2010 y diciembre 2011. Por último, la categoría de *empleado/obrero privado* es no significativa solamente en el estudio de Junio 2007 - Diciembre 2010.
- Dentro de las características geográficas: se puede observar que las *Zonas no delimitadas* resultan ser no significativas para los estudios de diciembre 2007 y diciembre 2008. Mientras que, *Quito* es no significativa para los estudios de diciembre 2011 y Junio 2007 - Diciembre 2007, *Ambato* resulta no significativa para los estudios de diciembre 2009 y Junio 2007 - Diciembre 2007. Finalmente, *Machala*, *Guayaquil* y el *Resto de la Costa* son no significativas en el estudio de Junio 2007 - Diciembre 2009.

4.1.1. Comparación con nuevos individuos

En este punto, es posible realizar la comparación de nuevos individuos con los modelos que tienen un mejor comportamiento en las pruebas contempladas anteriormente. Ciertamente, los modelos trabajados con los datos que incluyen información de diciembre de cada año desde el 2007 al 2011, en algunos casos, presentan una tasa de clasificación general mayor en relación a los modelos en los que se trabaja con datos agregados a manera de cascada. No obstante, el precio de tener variables importantes¹ dentro del modelo, que se tornan no significativas, como son características personales, laborales y geográficas, nos lleva a la conclusión de que el uso del modelo logit, por su inestabilidad, no es el más indicado para realizar la clasificación de nuevos individuos. Una vez efectuada la consideración de que pueden existir modelos inestables, para clasificar a nuevos individuos; consideremos el periodo de junio 2007 a junio 2010 y el periodo de junio 2007 a junio 2012. Debido a que, estos últimos tienen las si-

¹La expresión *Importante* hace referencia a que en las pruebas univariantes realizadas estas variables se tornaban significativas.

guientes características que, a nuestra opinión, lo convierten en un estudio más estable: la muestra es mayor en un 57 %, el porcentaje de clasificación y el R cuadrado de Nagelkerke son iguales y, finalmente, las variables incluidas en el estudio son significativas a excepción del nivel de instrucción.

A continuación, se realizará la clasificación de nuevos individuos con los dos estudios. Para comprobar qué estudio clasifica mejor a los nuevos individuos a evaluar, se tomará la base de individuos correspondiente al mes de diciembre de los años 2012, 2013 y 2014, respectivamente, para clasificarlos dentro de los dos estudios. Los resultados encontrados se pueden ver en las tablas 4.4, 4.5 y 4.6.

1. Base de individuos diciembre 2012

		Junio2007-Junio2010			Junio2007-Junio2012				
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	977	4212	5189	18,83 %	931	4258	5189	17,94 %
	Temporal	1370	3559	4929	72,21 %	1598	3331	4929	67,58 %
	Total	2347	7771	10118		2529	7589	10118	
% Global					44,83 %				42,12 %

Tabla 4.4: Tabla de Clasificación con nuevos individuos del año 2012. Fuente: elaboración propia a partir de la encuesta ENEMDU.

2. Base de individuos diciembre 2013

		Junio2007-Junio2010			Junio2007-Junio2012				
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	1297	4017	5314	24,41 %	1206	4108	5314	22,69 %
	Temporal	1563	3945	5508	71,62 %	1868	3640	5508	66,09 %
	Total	2860	7962	10822		3074	7748	10822	
% Global					48,44 %				44,78 %

Tabla 4.5: Tabla de Clasificación con nuevos individuos del año 2013. Fuente: elaboración propia a partir de la encuesta ENEMDU.

3. Base de individuos diciembre 2014

		Junio2007-Junio2010			Junio2007-Junio2012				
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	2154	6318	8472	25,42 %	1879	6593	8472	22,18 %
	Temporal	2820	5663	8483	66,76 %	3179	5304	8483	62,53 %
	Total	4974	11981	16955		5058	11897	16955	
% Global					46,10 %				42,37 %

Tabla 4.6: Tabla de Clasificación con nuevos individuos del año 2014. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En las tablas 4.4, 4.5 y 4.6, se puede observar que, para los nuevos individuos que se han usado como datos de verificación correspondientes a la información del mes de diciembre de los años: 2012, 2013, 2014 en el modelo junio2007-junio2010 (trabajo publicado) se obtienen los siguientes porcentajes de clasificación global y de sensibilidad:

- Individuos de verificación del 2012 obtienen el 44,83 % de clasificación global y 72,21 % en el porcentaje de sensibilidad.
- Individuos de verificación del 2013 obtienen el 48,44 % de clasificación global y 71,62 % en el porcentaje de sensibilidad.
- Individuos de verificación del 2014 obtienen el 46,10 % de clasificación global y 66,76 % en el porcentaje de sensibilidad.

En cambio, en el estudio junio2007-junio2012 alcanzan los siguientes porcentajes de clasificación global y de sensibilidad:

- Individuos de verificación del 2012 obtienen el 42,12 % de clasificación global y 67,58 % en el porcentaje de sensibilidad.
- Individuos de verificación del 2013 obtienen el 44,78 % de clasificación global y 66,09 % en el porcentaje de sensibilidad.
- Individuos de verificación del 2014 obtienen el 42,37 % de clasificación global y 62,53 % en el porcentaje de sensibilidad.

En este sentido, se puede concluir que la clasificación de los nuevos individuos de verificación usando el estudio junio 2007- junio 2010 (trabajo publicado)

obtiene una mayor tasa de clasificación global y un mayor porcentaje de sensibilidad. En los tres casos, aproximadamente un 4% más que lo que se alcanza con el estudio junio 2007 -junio 2012.

4.2. Modelo de Máquinas de Vectores Soporte (MVS)

Esta sección tiene como objetivo comparar los resultados que se obtuvieron con la regresión logística. Para ello, se considerarán dos condiciones iniciales, las cuales tienen como fin permitir la comparación o contrastes entre los dos modelos. Las condiciones son:

- i Se usará la misma muestra de datos.
- ii Se usará las mismas variables dependientes y la misma variable independiente.

Por su parte, la metodología utilizada para la estimación de MVS se lleva a cabo en los siguientes pasos:

- a) Introducción de la función costo natural C , ecuación (2.21), que permite cierto margen de error. Representa un balance entre el tamaño del margen y el error de entrenamiento [38].
- b) Introducción de la función kernel seleccionada, que consiste en asignar a los datos de entrada un espacio de características de mayor dimensión [41].
- c) Selección de la combinación de parámetros que genera menor error a través de la validación cruzada de K iteraciones o K -fold cross-validation (Ver Anexo B).
- d) Se construye o programa la máquina de vectores soporte para clasificación con los parámetros seleccionados.
- e) Se clasifica a los individuos seleccionados utilizando las bases elegidas y la máquina de vectores soporte programada.

4.2.1. Estimación de la MVS

La estimación de la máquina de vectores soporte toma en cuenta los cuatro puntos que se detallaron anteriormente. Respecto al numeral ii, en esta tesis, se usarán las funciones kernels que son recomendadas en [33] y [38]; especialmente las siguientes:

- Kernel lineal: $K(x, x') = \langle x, x' \rangle$.
- Kernel polinomial de grado d : $K_d(x, x') = [\gamma \langle x, x' \rangle + \gamma]^d$.
- Kernel radial gaussiana $K(x, x') = \exp(-\gamma \|x - x'\|^2)$, son: $\gamma > 0$
- Kernel sigmoidal: $K(x, x') = \tanh(\gamma \langle x, x' \rangle + \tau)$

Respecto a la introducción (selección de parámetros C y γ) y la estimación de MVS, en la figura 4.1 se presenta un esquema que permitirá conocer el procedimiento para obtener la estimación.

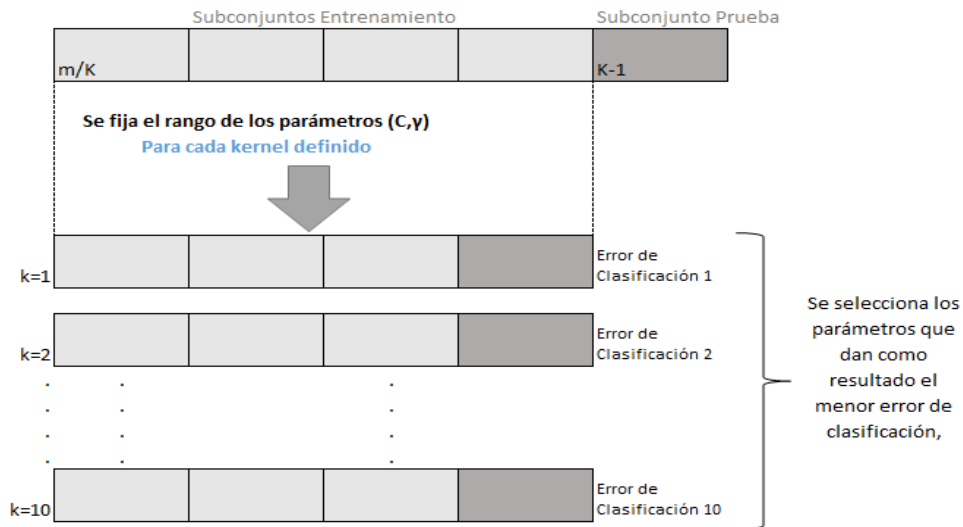
Proceso de elección de parámetros

Para el proceso de elección de parámetros C y γ , se ha tomado la decisión de usar los datos correspondientes al levantamiento realizado en el mes de diciembre de los años 2007, 2008, 2009, 2010 y 2011, de manera individual y no agregada, puesto que al utilizar la información de manera agregada no se lograba obtener convergencia con el software elegido; esto obligó a trabajar con bases más pequeñas, las que permitieran llegar a obtener convergencia y unos parámetros C y γ a ser elegidos, como es lo recomendado en [33]. El parámetro de la función kernel es representado por γ , en tanto que el parámetro de penalización o regularización es C .

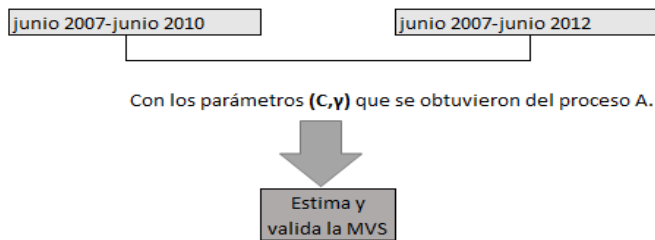
Dado que no existe una manera teórica para la estimación del parámetro de regularización [12], se usará el siguiente proceso para la estimación de los parámetros C y γ :

1. El rango de intervalo de prueba que se usará para C será $[10^0, 10^2]$ con un tamaño de paso constante de 1 en el exponente; es decir, se tomarán los valores: 1, 10 y 100.

A) Para cada mes de diciembre del 2007 al 2011 se ejecuta este proceso



B) Con la información agregada junio 2007-junio 2010 y junio 2007 - junio 2012 se estima la MVS, con los parámetros que se obtienen del proceso A.



C) Con la información del mes de diciembre de los años 2012, 2013, 2014 se realiza la validación.

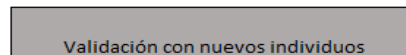


Figura 4.1: Esquema de estimación de la MVS. Fuente: elaboración propia.

2. El rango de intervalo que se tomará para γ será $[10^{-6}, 10^{-3}]$ con un tamaño de paso 1 en el exponente; en otras palabras, se tomarán los valores: 0,000001; 0,00001; 0,0001 y 0,001.
3. El método que se usará para realizar la estimación es K-Fold Cros Validation con $K = 10$ (Ver Anexo B) para cada función kernel. Para ello, se escoge los parámetros C y γ que son los que arrojan el menor error de clasificación.

De esta manera, los resultados que se obtienen son:

Periodo	Radial Gaussian			Lineal			Polinomial(2)			Sigmoidal		
	C	γ	Error	C	γ	Error	C	γ	Error	C	γ	Error
dic-07	100	0,001	0,2338	10	0,000001	0,2339	100	0,001	0,4686	10	0,001	0,2334
dic-08	100	0,001	0,2075	1	0,000001	0,2105	100	0,001	0,4521	10	0,001	0,2118
dic-09	100	0,001	0,2129	1	0,000001	0,2247	100	0,001	0,4451	100	0,001	0,2235
dic-10	100	0,001	0,2214	10	0,000001	0,2409	100	0,001	0,4580	1	0,001	0,2371
dic-11	100	0,001	0,2348	1	0,000001	0,2502	100	0,001	0,4810	10	0,00001	0,2447

Tabla 4.7: Elección de Parámetros C y γ . Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la tabla 4.7 se puede observar el valor de C y γ con su respectivo error al obtener los parámetros para cada tipo de función kernel por cada año, respectivamente.

Estimación de la MVS

Los datos correspondientes a la unión en cascada que se señala en la sección 4.1, desde el mes de junio 2007 a junio 2010 y desde el junio 2007 a junio 2012 serán utilizados para la estimación de la máquina de vectores soporte para clasificación; esto se realizará con los parámetros seleccionados, de acuerdo al proceso que se detalló en el numeral anterior.

Dado que para cada periodo de análisis se evaluó cada una de las funciones kernels, y se obtuvo el correspondiente valor de C y γ , para la estimación de la MVS con la información agregada se usarán los parámetros que se alcanzarán, de la siguiente manera:

Kernel	C	Gamma	Error
Radial Gaussian	100	0,001	0,2221
Lineal	10	0,000001	0,2374
	1	0,000001	0,2285
Polinomial	100	0,001	0,4610
Sigmoidal	10	0,001	0,2226
	100	0,001	0,2235
	1	0,001	0,2371
	10	0,000001	0,2447

Tabla 4.8: Parámetros C y γ para modelos agregados. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Los valores que se muestran en la tabla 4.8 serán los parámetros con los que se realizará la estimación de los distintos modelos agregados. Los resultados alcanzados son los siguientes:

1. Para la base unificada desde junio 2007 a junio 2010 se tiene:

a) Kernel Radial Gausiann:

Tabla de clasificación					
Observado		Pronosticado			Porcentaje correcto
		Y			
		Contrato Indefinido	Contrato Temporal	Total General	
Y	Contrato Indefinido	21 644	9 466	31 110	69,57
	Contrato Temporal	5 916	29 629	35 545	83,36
	Total General	27 560	39 095	66 655	
Porcentaje global					76,92

Tabla 4.9: Tabla de Clasificación del modelo jun07jun10 con kernel Radial $C = 100$ y $\gamma = 0,001$. Fuente: elaboración propia a partir de la encuesta ENEMDU.

b) Kernel Lineal.

Tabla de clasificación $C = 10$ $\gamma = 0,000001$					
Observado		Pronosticado			Porcentaje correcto
		Y			
		Contrato Indefinido	Contrato Temporal	Total General	
Y	Contrato Indefinido	23 362	7 748	31 110	75,09
	Contrato Temporal	7 465	28 080	35 545	79,00
	Total General	30 827	35 828	66 655	
Porcentaje global					77,18
Tabla de clasificación $C = 1$ $\gamma = 0,000001$					
Observado		Pronosticado			Porcentaje correcto
		Y			
		Contrato Indefinido	Contrato Temporal	Total General	
Y	Contrato Indefinido	23 373	7 737	31 110	75,13
	Contrato Temporal	7 474	28 071	35 545	78,97
	Total General	30 827	35 828	66 655	
Porcentaje global					77,18

Tabla 4.10: Tabla de clasificación del periodo jun07jun10 con kernel Lineal.

c) Kernel Polinomial de grado 2.

Tabla de clasificación					
Observado		Pronosticado			Porcentaje correcto
		Y			
		Contrato Indefinido	Contrato Temporal	Total General	
Y	Contrato Indefinido	21 877	9 233	31 110	70,32
	Contrato Temporal	5 904	29 641	35 545	83,39
	Total General	27 781	38 874	66 655	
Porcentaje global					77,29

Tabla 4.11: Tabla de Clasificación del periodo jun07jun10 con kernel Polinomial $C = 100$ y $\gamma = 0,001$. Fuente: elaboración propia a partir de la encuesta ENEMDU.

d) Kernel Sigmoidal

Tabla de clasificación $C = 10 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	23 336	7 774	31 110	75,01	
	Contrato Temporal	7 464	28081	35 545	79,00	
	Total General	30 800	35 855	66 655		
Porcentaje global					77,14	
Tabla de clasificación $C = 100 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	23 339	7 771	31 110	75,02	
	Contrato Temporal	7 744	27 801	35 545	78,21	
	Total General	31 083	35 572	66 655		
Porcentaje global					76,72	
Tabla de clasificación $C = 1 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	23 858	7 252	31 110	76,69	
	Contrato Temporal	8 087	27458	35 545	77,25	
	Total General	31 945	34 710	66 655		
Porcentaje global					76,99	
Tabla de clasificación $C = 10 \gamma = 0,000001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	22 564	8 546	31 110	72,53	
	Contrato Temporal	6 917	28 628	35 545	80,54	
	Total General	29 481	37 174	66 655		
Porcentaje global					76,80	

Tabla 4.12: Tabla de clasificación del periodo jun07jun10 con kernel Sigmoidal.

2. Para la base unificada desde junio 2007 a junio 2012 se tiene:

a) Kernel Radial Gausiann:

Tabla de clasificación						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	34 654	14 788	49 442	70,09	
	Contrato Temporal	9 575	45 587	55 162	82,64	
	Total General	49 442 560	60 375	104 604		
Porcentaje global					76,71	

Tabla 4.13: Tabla de clasificación del periodo jun07jun12 con kernel Radial $C = 100$ y $\gamma = 0,001$. Fuente: elaboración propia a partir de la encuesta ENEMDU.

b) Kernel Lineal.

Tabla de clasificación $C = 10$ $\gamma = 0,000001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	38 713	10 729	49 442	78,30	
	Contrato Temporal	16 709	38 453	55 162	69,71	
	Total General	55 422	49 182	104 604		
Porcentaje global					73,77	

Tabla de clasificación $C = 1$ $\gamma = 0,000001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	39 522	9 920	49 442	79,94	
	Contrato Temporal	14 937	40 225	55 162	72,92	
	Total General	55 422	49 182	104 604		
Porcentaje global					76,24	

Tabla 4.14: Tabla de clasificación del periodo jun07jun12 con kernel Lineal.

c) Kernel Polinomial de grado 2.

Tabla de clasificación						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	34 975	14 467	49 442	70,74	
	Contrato Temporal	9 403	45 759	55 162	82,95	
	Total General	44 378	60 226	104 604		
Porcentaje global					77,18	

Tabla 4.15: Tabla de clasificación del periodo jun07jun12 con kernel Polinomial $C = 100$ y $\gamma = 0,001$. Fuente: elaboración propia a partir de la encuesta ENEMDU.

d) Kernel Sigmoidal

Tabla de clasificación $C = 10 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	39 553	9 889	49 442	80,00	
	Contrato Temporal	15 013	40 149	55 162	72,78	
	Total General	54 566	50 038	104 604		
Porcentaje global					76,19	
Tabla de clasificación $C = 100 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	39 000	10 442	49 442	78,88	
	Contrato Temporal	14 807	40 355	55 162	73,16	
	Total General	53 807	50 797	104 604		
Porcentaje global					75,88	
Tabla de clasificación $C = 1 \gamma = 0,001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	39 770	9 672	49 442	80,44	
	Contrato Temporal	15 560	39 602	55 162	71,79	
	Total General	55 330	49 274	104 604		
Porcentaje global					75,88	
Tabla de clasificación $C = 10 \gamma = 0,000001$						
Observado		Pronosticado				Porcentaje correcto
		Y				
		Contrato Indefinido	Contrato Temporal	Total General		
Y	Contrato Indefinido	39 396	10 046	49 442	79,68	
	Contrato Temporal	14 719	40 443	55 162	73,32	
	Total General	54 115	50 489	104 604		
Porcentaje global					76,32	

Tabla 4.16: Tabla de clasificación del periodo jun07jun12 con kernel Sigmoidal.

En las tablas 4.9, 4.10, 4.11, 4.12, 4.13, 4.14 4.15 y 4.16, se puede observar que se obtiene un porcentaje global de clasificación que varía aproximadamente entre el 76% y el 77%. Es el modelo con función kernel polinomial de grado 2, $C = 100$ y $\gamma = 0,001$ el que consigue el mayor porcentaje global de clasificación tanto para el modelo correspondiente a junio 2007-junio 2010 y el de junio 2007-junio 2012; este porcentaje es de 77,29% y de 77,18% respectivamente.

Comprobación con nuevos individuos

Los datos pertenecientes al mes de diciembre de los años 2012, 2013 y 2014 corresponden a los datos que serán utilizados en la prueba final de validación.

Es importante evaluar cómo clasifica el modelo a nuevos individuos; a causa de esto, se presentan los resultados que se obtuvieron con cada uno de los modelos desarrollados:

1. Diciembre 2012

a) Kernel Radial Gausiann:

Junio2007-Junio2010					Junio2007-Junio2012				
Observado	Tabla clasif.con nuevos individuos				Tabla clasif.con nuevos individuos				
Pronosticado									
Y				% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	3959	1230	5189	76,30 %	3690	1499	5189	71,11 %
	Temporal	1686	3243	4929	65,79 %	973	3956	4929	80,26 %
	Total	5645	4473	10118		4663	5455	10118	
% Global				71,18 %				75,57 %	

Tabla 4.17: Tabla de clasificación de nuevos individuos diciembre 2012. Fuente: elaboración propia a partir de la encuesta ENEMDU.

b) Kernel Lineal.

Junio2007-Junio2010					Junio2007-Junio2012				
Observado	Tabla clasif.con nuevos individuos				Tabla clasif.con nuevos individuos				
Pronosticado									
Y				% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4178	1011	5189	80,52 %	4345	844	5189	83,73 %
	Temporal	1547	3382	4929	68,61 %	1571	3358	4929	68,13 %
	Total	5725	4393	10118		5916	4202	10118	
% Global				74,72 %				76,13 %	
Observado	Tabla clasif.con nuevos individuos				Tabla clasif.con nuevos individuos				
Pronosticado									
Y				% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4184	1005	5189	80,63 %	4329	860	5189	83,43 %
	Temporal	1552	3377	4929	68,51 %	1501	3428	4929	69,55 %
	Total	5736	4382	10118		5830	4288	10118	
% Global				74,73 %				76,67 %	

Tabla 4.18: Tabla de clasificación de nuevos individuos kernel Lineal.

c) Kernel Polinomial de grado 2.

Junio2007-Junio2010					Junio2007-Junio2012				
Observado	Tabla clasif.con nuevos individuos				Tabla clasif.con nuevos individuos				
Pronosticado									
Y				% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	3856	1333	5189	74,31 %	3778	1411	5189	72,81 %
	Temporal	1361	3568	4929	72,39 %	997	3932	4929	79,77 %
	Total	5217	4901	10118		4775	5343	10118	
% Global				73,37 %				76,20 %	

Tabla 4.19: Tabla de clasificación con nuevos individuos con kernel Polinomial. Fuente: elaboración propia a partir de la encuesta ENEMDU.

d) Kernel Sigmoidal

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4183	1006	5189	80,61 %	4330	859	5189	83,45 %
	Temporal	1547	3382	4929	68,61 %	1502	3427	4929	69,53 %
	Total	5730	4388	10118		5832	4286	10118	
% Global				74,77 %	76,67 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4171	1018	5189	80,38 %	4299	890	5189	82,85 %
	Temporal	1582	3347	4929	67,90 %	1500	3429	4929	69,57 %
	Total	5753	4365	10118		5799	4319	10118	
% Global				74,30 %	76,38 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4217	972	5189	81,27 %	4369	820	5189	84,20 %
	Temporal	1554	3375	4929	68,47 %	1535	3394	4929	68,86 %
	Total	5771	4347	10118		5904	4214	10118	
% Global				75,03 %	76,72 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4121	1068	5189	79,42 %	4312	877	5189	83,10 %
	Temporal	1562	3367	4929	68,31 %	1492	3437	4929	69,73 %
	Total	5683	4435	10118		5804	4314	10118	
% Global				74,01 %	76,59 %				

Tabla 4.20: Tabla de clasificación con nuevos individuos con kernel Sigmoidal.

2. Diciembre 2013:

a) Kernel Radial Gausiann:

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4042	1272	5314	76,06 %	3910	1404	5314	73,58 %
	Temporal	1874	3634	5508	65,98 %	1229	4279	5508	77,69 %
	Total	5916	4906	10822		5139	5683	10822	
% Global				70,93 %	75,67 %				

Tabla 4.21: Tabla de clasificación con nuevos individuos con kernel Radial. Fuente: elaboración propia a partir de la encuesta ENEMDU.

b) Kernel Lineal.

		Junio2007-Junio2010			Junio2007-Junio2012				
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	4343	971	5314	81,73 %	4424	890	5314	83,25 %
	Temporal	1806	3702	5508	67,21 %	1734	3774	5508	68,52 %
	Total	6149	4673	10822		6158	4664	10822	
% Global					74,34 %				75,75 %
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	4347	967	5314	81,80 %	4486	828	5314	84,42 %
	Temporal	1812	3696	5508	67,10 %	1724	3784	5508	68,70 %
	Total	6159	4663	10822		6210	4612	10822	
% Global					74,32 %				76,42 %

Tabla 4.22: Tabla de clasificación con nuevos individuos con kernel Lineal.

c) Kernel Polinomial de grado 2.

		Junio2007-Junio2010			Junio2007-Junio2012				
Observado		Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos				
		Pronosticado							
		Y			% Correcto	Y			% Correcto
		Indefinido	Temporal	Total		Indefinido	Temporal	Total	
Y	Indefinido	3924	1390	5314	73,84 %	3963	1351	5314	74,58 %
	Temporal	1486	4022	5508	73,02 %	1183	4325	5508	78,52 %
	Total	5410	5412	10822		5146	5676	10822	
% Global					73,42 %				76,58 %

Tabla 4.23: Tabla de clasificación con nuevos individuos con kernel Polinomial. Fuente: elaboración propia a partir de la encuesta ENEMDU.

d) Kernel Sigmoidal

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
	Pronosticado								
	Y			% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4344	970	5314	81,75 %	4488	826	5314	84,46 %
	Temporal	1802	3706	5508	67,28 %	1719	3789	5508	68,79 %
	Total	6146	4676	10822		6207	4615	10822	
% Global				74,39 %	76,48 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
	Pronosticado								
	Y			% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4349	965	5314	81,84 %	4456	858	5314	83,85 %
	Temporal	1864	3644	5508	66,16 %	1731	3777	5508	68,57 %
	Total	6213	4609	10822		6187	4635	10822	
% Global				73,86 %	76,08 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
	Pronosticado								
	Y			% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4378	936	5314	82,39 %	4492	822	5314	84,53 %
	Temporal	1762	3746	5508	68,01 %	1738	3770	5508	68,45 %
	Total	6140	4682	10822		6230	4592	10822	
% Global				75,07 %	76,34 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
	Pronosticado								
	Y			% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	4240	1074	5314	79,79 %	4465	849	5314	84,02 %
	Temporal	1714	3794	5508	68,88 %	1655	3853	5508	69,95 %
	Total	5954	4868	10822		6120	4702	10822	
% Global				74,24 %	76,86 %				

Tabla 4.24: Tabla de clasificación con nuevos individuos con kernel Sigmoidal.

e) Diciembre 2014:

1) Kernel Radial Gausiann:

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
	Pronosticado								
	Y			% Correcto	Y			% Correcto	
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6403	2069	8472	75,58 %	5936	2536	8472	70,07 %
	Temporal	2992	5491	8483	64,73 %	2101	6382	8483	75,23 %
	Total	9395	7560	16955		8037	8918	16955	
% Global				70,15 %	72,65 %				

Tabla 4.25: Tabla de Clasificación con nuevos individuos con kernel Radial. Fuente: elaboración propia a partir de la encuesta ENEMDU.

2) Kernel Lineal.

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6716	1756	8472	79,27 %	6956	1516	8472	82,11 %
	Temporal	2827	5656	8483	66,67 %	2636	5847	8483	68,93 %
	Total	9543	7412	16955		9592	7363	16955	
% Global				72,97 %	75,51 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6723	1749	8472	79,36 %	6979	1493	8472	82,38 %
	Temporal	2832	5651	8483	66,62 %	2683	5800	8483	68,37 %
	Total	9555	7400	16955		9662	7293	16955	
% Global				72,98 %	75,37 %				

Tabla 4.26: Tabla de clasificación con nuevos individuos con kernel Lineal.

3) Kernel Polinomial de grado 2.

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6115	2357	8472	72,18 %	6044	2428	8472	71,34 %
	Temporal	2370	6113	8483	72,06 %	1997	6486	8483	76,46 %
	Total	8485	8470	16955		8041	8914	16955	
% Global				72,12 %	73,90 %				

Tabla 4.27: Tabla de Clasificación con nuevos individuos con kernel Polinomial. Fuente: elaboración propia a partir de la encuesta ENEMDU.

4) Kernel Sigmoidal

Junio2007-Junio2010				Junio2007-Junio2012					
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6710	1762	8472	79,20 %	6983	1489	8472	82,42 %
	Temporal	2792	5691	8483	67,09 %	2684	5799	8483	68,36 %
	Total	9502	7453	16955		9667	7288	16955	
% Global				73,14 %	75,39 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6736	1736	8472	79,51 %	6941	1531	8472	81,93 %
	Temporal	2876	5607	8483	66,10 %	2712	5771	8483	68,03 %
	Total	9612	7343	16955		9653	7302	16955	
% Global				72,80 %	74,97 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6796	1676	8472	80,22 %	7007	1465	8472	82,71 %
	Temporal	2721	5762	8483	67,92 %	2680	5803	8483	68,41 %
	Total	9517	7438	16955		9687	7268	16955	
% Global				74,07 %	75,55 %				
Observado	Tabla clasif.con nuevos individuos			Tabla clasif.con nuevos individuos					
Pronosticado									
Y			% Correcto	Y			% Correcto		
	Indefinido	Temporal	Total		Indefinido	Temporal	Total		
Y	Indefinido	6605	1867	8472	77,96 %	6963	1509	8472	82,19 %
	Temporal	2633	5850	8483	68,96 %	2615	5868	8483	69,17 %
	Total	9238	7717	16955		9578	7377	16955	
% Global				73,46 %	75,68 %				

Tabla 4.28: Tabla de clasificación con nuevos individuos con kernel Sigmoidal.

A partir de la tabla 4.17 hasta la tabla 4.28 se puede observar que tanto para junio 2007- junio 2010 y junio 2007-junio 2012 el mayor porcentaje de clasificación global se obtiene cuando se usa la función kernel Sigmoidal con $C = 1$ y $\gamma = 0,001$. La única excepción está presente en la validación con nuevos individuos de los años 2013 y 2014 en el modelo de junio 2007 - junio 2012 que tiene la mayor tasa de clasificación cuando se usa la función kernel Sigmoidal con $C = 10$ y $\gamma = 0,00001$.

Es preciso notar que, a pesar de tener una mayor clasificación global con la función kernel Sigmoidal que con la función kernel polinomial de grado 2, con esta última se tiene una sensibilidad mayor, en relación al resto de modelos evaluados.

Por tal razón, el modelo a considerarse en MVS es el estimado a partir de la función polinomial de grado 2, $C = 100$ y $\gamma = 0,001$.

Capítulo 5

Conclusiones y Recomendaciones

En este capítulo, se resume el estudio realizado en la presente tesis de grado, analizando los resultados obtenidos y el cumplimiento de los objetivos planteados en el capítulo 1. Así mismo, se plantean recomendaciones en torno a futuros trabajos relacionados con este tema que se consideran de especial interés en el ámbito académico.

5.1. Conclusiones

La presente tesis ha tenido como uno de sus objetivos estudiar las principales características económicas que influyen para la contratación temporal y determinar en qué medida afectan al momento de conseguir un empleo. De esta investigación, se puede concluir lo siguiente:

- A partir de datos de la Encuesta de Empleo, Desempleo y Subempleo del Instituto Nacional de Estadísticas y Censos del Ecuador y de la aplicación de un modelo de regresión logística, usando la desagregación de las características personales, laborales y geográficas de los individuos, este trabajo ayuda a comprender la realidad laboral ecuatoriana, pues permite detectar las diferencias en el acceso al empleo temporal según la rama de actividad económica.
- Los resultados obtenidos han mostrado que la temporalidad en Ecuador no solo depende de la estructura productiva existente, sino que se trata de un

fenómeno asociado a las características de los trabajadores y a nuestra cultura empresarial. Resalta la influencia que tiene el nivel de instrucción, el tiempo de trabajo, el tipo de jornada laboral, la capacitación al momento de obtener un contrato temporal, y el género de quien aplica para el empleo. Los hombres presentan una mayor propensión de tener un contrato temporal. Esto podría explicarse por el hecho de que existen tareas sencillas y rudimentarias, como las que exigen los cultivos agrícolas, la cría de ganado, la pesca, la caza, la industria y construcción que requieran la utilización de herramientas manuales y, a menudo, un esfuerzo físico considerable.

- Se observa que el tipo de ocupación es una variable influyente sobre la determinación del tipo de contrato que la persona tenga o pueda llegar a obtener. Se evidencia, así, que si una persona tiene una ocupación de jornalero o peón, es esta una condición que afecta fuertemente a la probabilidad de tener un contrato temporal. Esto lleva a un estancamiento de la situación laboral de las personas que se desenvuelven dentro de esta ocupación particular, logrando que no puedan tener un nivel salarial ascendente; lo que afecta directamente el nivel económico de las familias.
- En referencia a las ciudades autorepresentadas y regiones, se obtiene que en la Región Costa, particularmente en las ciudades de Machala y Guayaquil, se concentra la modalidad de los contratos temporales. Una posible explicación se encontraría en las actividades, como: agricultura, ganadería, caza y silvicultura; industrias manufactureras; hoteles y restaurantes; transporte; almacenamiento y comunicaciones. En las áreas mencionadas, se pueden presentar contratos eventuales, por obra o por horas; y generalmente son actividades que se desarrollan por temporadas, es decir, en ciertos períodos del año surge la necesidad de incrementar el contrato temporal.

Por otra parte, el segundo de los objetivos hace referencia a analizar y comparar una técnica clásica, la regresión logística, versus una contemporánea, máquina de vectores soporte, con la intención de averiguar sus alcances y limitaciones; a partir de ello, se puede concluir lo siguiente:

- Enfocándose en el procedimiento que se debe llevar a cabo y en el costo

computacional de las dos técnicas evaluadas, se puede destacar que la Regresión Logística es un método que se caracteriza por su facilidad de implementación, en tanto que la Máquina de Vectores Soporte requiere de un mayor nivel computacional.

- Respecto a la interpretación de los parámetros que se obtienen en los modelos la tabla de clasificación de individuos es similar y es el factor común con el que se realizó la comparación. Los valores individuales por cada variable independiente que se los interpreta como los Odds-ratio mediante la Regresión Logística, no pueden ser obtenidos a través de la Máquina de Vectores Soporte.
- De acuerdo al porcentaje global de clasificación que se obtuvo mediante el modelo de Regresión Logística y la Máquina de Vectores Soporte con kernel Polinomial de grado 2, $C = 100$ y $\gamma = 0,01$, los valores oscilan al rededor del 77% para las dos técnicas.
- Se puede notar una mejora importante al momento de clasificar nuevos individuos. El porcentaje de clasificación global para nuevos individuos con la técnica Máquina de Vectores Soporte es mayor aproximadamente en un 30% al que se obtiene con la técnica de Regresión Logística.

5.2. Recomendaciones

Respecto a la aplicación de la máquina de vectores soporte, se destaca que la mayor dificultad del método recae en la calibración o proceso de estimación de algunos parámetros asociados al Kernel o Núcleo utilizado y al parámetro de penalización del modelo. Por esta razón, un estudio que permita construir una técnica analítica o heurística para la determinación de estos parámetros a un menor costo computacional, sería un gran aporte para desarrollar esta metodología.

Referencias

- [1] Acosta A., (2005), "Breve historia Económica del Ecuador", Corporación Editorial Nacional, Quito, págs. 202-230.
- [2] Aguayo M., (2007), "Cómo hacer una Regresión Logística con SPSS paso a paso. (I)", Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702012, págs. 1-16.
- [3] Aguayo M., y Lore E., (2007), "Cómo hacer una Regresión Logística binaria paso a paso II análisis multivariante", Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702013, págs. 1-35.
- [4] Aguayo M., y Monge L., (2007), "Confusión e iteración (2): su abordaje en el análisis multivariante", Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702011, pág. 5.
- [5] Aguiar V., (2007), "El mercado Laboral Ecuatoriano: Propuesta de una reforma", Edición electrónica. Texto completo en www.eumed.net/libros/2007a/240/, ISBN: 978-84-690-5504-5, págs. 5- 25.
- [6] Alejandro A., Pérez J., Kizys R. y Manzanedo L., (2002), "Regresión Logística Binaria", Universidad Oberta de Catalunya, Barcelona, págs. 5-16.
- [7] Banco Interamericano de Desarrollo, (2011).
- [8] Becker G., (1962), "Investment in human capital: A Theoretical analysis", *Journal of Political Economy*, Volumen 70, págs. 9-49.
- [9] Boser B., Guyon I. y Vapnik V., (1992), "A Training Algorithm for Optimal Margin Classifiers. En: Proceedings of the fifth Annual Workshop on Computational Learning Theory", ACM Press, Nueva York, Nueva York, Estados Unidos, págs. 144-152.

- [10] Caparros A., y Navarro L., (2008), "Temporalidad, segmentación laboral y actividad productiva: ¿existen diferencias regionales?", Estadística Española, Vol. 50, Nº. 168, págs. 206-226.
- [11] Caparros A., Navarro L., y Rueda M., (2004), "Efectos de la temporalidad sobre la formación recibida durante el empleo", Universidad de Málaga, Cuadernos de Economía. Vol 27,051-074, págs.52-73.
- [12] Carmona E., (2014) "Tutorial sobre Máquinas de Vectores Soporte (SVM)", Dpto. de Inteligencia Artificial, ETS de Ingeniería Informática, Universidad Nacional de Educación a Distancia (UNED), Madrid, España.
- [13] Codificación 17, (2005), "Codificación del Código del Trabajo," Registro Oficial Suplemento 167, Diciembre, título I, capítulo I.
- [14] Chávez Y. y Medina P., (2012), "Determinantes de la Temporalidad en el Mercado laboral ecuatoriano", Revista de Métodos Cuantitativos para la Economía y la Empresa, ISSN: 1886-516X. D.L: SEP-2927-06, págs.24-53.
- [15] Cuadras C., (2010), "Nuevos métodos de análisis multivariante", CMC Editions Manacor 3008023 Barcelona, Spain, ISSN: 1886-516X. D.L: SEP-2927-06, págs.207-222.
- [16] Domínguez E. y Aldana D., (2001), "Regresión Logística. Un ejemplo de su uso en endocrinología", Rev Cubana Endocrinol, Vol. 12, págs. 58-64.
- [17] Gallardo C. y Capa H., (2008), "Nueva Metodología y Calculo de Indicadores de la Encuesta de Empleo, Desempleo y Subempleo", Revista Ecuatoriana de Estadística, Instituto Nacional de Estadísticas y Censos del Ecuador, págs. 51-53.
- [18] Gamero C., (1999), "Satisfacción Laboral y tipo de contrato en España", Investigative Radiology 34, Vol. 10, 636 - 642, págs. 421-429.
- [19] González L., (2002), "Análisis discriminante utilizando máquinas núcleos de vectores soporte función núcleo similitud", Universidad de Sevilla, págs. 1-393.
- [20] Hachuel L., Boggio G., Wojdyla D. y Servy E., (2005), "Interpretación y comparación de modelos de regresión logística para el estudio de la desocupa-

ción”, Décimas Jornadas Investigaciones en la Facultad de Ciencias Económicas y Estadística, Instituto de Investigaciones Teóricas y Aplicadas, págs. 1-12.

- [21] Hamel L., (2009), “Knowledge Discovery with Support Vector Machine”, John Wiley y Sons, págs. 91-130.
- [22] Estrada B. y Mera C., (2005), “ Support Vector Machine”, Universidad Nacional de Colombia, págs. 1-60.
- [23] Instituto Nacional de Estadísticas y Censos del Ecuador, (2010), “Estadísticas de Empleo, Mercado Laboral”. Disponible en: www.inec.gob.ec/web/guest/ecu_est/est_soc/enc_hog/enemdu.
- [24] Instituto Nacional de Estadísticas y Censos del Ecuador, (2007-2010) “Estadísticas de Empleo, Mercado Laboral”. Disponible en: www.ecuadorencifras.com/cifras--inec/bancInf.html.
- [25] Instituto Nacional de Estadísticas y Censos del Ecuador, (2010), “Manual del Encuestador”, Estadística de Hogares, págs.54-88.
- [26] Jovanovic B., (1979), “Job Matching and the Theory of Turnover”, The Journal of Political Economy, Vol. 87, No. 5, Part 1, págs. 972-990.
- [27] Ley No. 100, (2003), “Código de la niñez y adolescencia”, Registro Oficial 737, Enero, págs. 24-26.
- [28] Ley de Maquilas y contratación laboral a tiempo parcial , (1990), “Código del Trabajo”, Registro Oficial 493, Agosto, Título II.
- [29] Ley para la Transformación Económica del Ecuador “Trole 1”, (2000), “Código del Trabajo”, Registro Oficial 33, Marzo, Capítulo XII.
- [30] Loh E., (1994), “Employment probation as a sorting mechanism”, Industrial and Labour Relations Review, Vol. 47, No. 3, págs. 471-486.
- [31] Mandato Constituyente No. 8, (2008), “Eliminación y prohibición de la tercerización, Intermediación Laboral, Contratación Laboral por horas y cualquier forma de precarización de las relaciones de trabajo”, Suplemento del Registro Oficial 330,6-V-2008.

- [32] E. Medina, (2003), "Modelos de Elección Discreta", Estadística de Hogares, www.eva.medinaam.es, págs. 9-21.
- [33] Moreno J. y Melo L.,(2011) " Pronóstico de incumplimientos de pago mediante máquinas de vectores de soporte: una aproximación inicial a la gestión del riesgo de crédito", Banco de la República, Borradores de Economía, no. 677, págs. 1-31.
- [34] Peña D., (2010), "Regresión y Diseño de Experimentos", Alianza Editorial, Spain, ISBN: 978-84-206-9389-7, págs.639-659.
- [35] —, "Regresión Logística Binaria," http://www.uoc.edu/in3/emath/docs/T10_Reg_Logistica.pdf
- [36] —, "Regresión Logística", http://www.seqc.es/es/Enlaces/6/Revistas_y_publicaciones
- [37] R Core Team,(2014), " R: A Language and Environment for Statistical Computing", R Foundation for Statistical Computing, Vienna, Austria. Disponible en: <http://www.R-project.org/>.
- [38] Robotham H., Bosh P., Castillo J. y Tapia I.,(2012) " Clasificación acústica de anchoveta (*Engraulis ringens*) y sardina común (*Strangomera bentincki*) mediante máquinas de vectores soporte en la zona centro-sur de Chile: efecto de la calibración de los parámetros en la matriz de confusión", Lat. Am. J. Aquat. Res., vol.40, no.1, págs. 90-101, ISSN 0718-560X.
- [39] Rodríguez O.,(2012) "Validación cruzada (cross-validation y remuestreo bootstrapping)", Programa Iberoamericano de Formación en Minería de Datos. Disponible en: <http://rpubs.com/orodriguez/13317>.
- [40] Salazar D., (2012), "Comparación de Máquinas de Soporte Vectorial vs. Regresión Logística. ¿Cuál es más recomendable para discriminar?", Massachusetts Institute of Technology, ISBN:0-262-19475-9, págs. 1-55.
- [41] Schölkopf B. y Smola A., (2002), "Learning with Kernels", Massachusetts Institute of Technology, ISBN:0-262-19475-9, págs. 1-55.
- [42] Steinwart I. y Christman A., (2008), "Support Vector Machines", Springer Scienci, ISBN:978-0-387-77241-7, págs. 1-19.

- [43] Solera R., (2011), "Máquinas de Vectores Soporte para reconocimiento robusto de habla", Universidad Carlos III Madrid, págs. 25-62.
- [44] Vapnik V., (1995), "The Nature of Statistical Learning Theory", Springer-Verlag, págs. 19-48.
- [45] Vapnik V., (1998), " Statistical Learning Theory", John Wiley & Sons, págs. 20-53.

Anexos

A. Análisis Descriptivo

Características Personales

En las tablas A.1, A.2, A.3 y A.4 se encuentra la información desglosada en lo que respecta a las variables que hacen referencia a las características personales que discriminan de mejor manera la proporción de contratos temporales vs. contratos indefinidos. Notándose que el contrato temporal se concentra en mayor medida en los hombres en relación a las mujeres, que presentan una menor propensión en la concentración del contrato temporal con una concentración del 18,59%, recalcando de igual manera que en la mayor parte del universo analizado está conformado por hombres.

Se puede evidenciar que el rango de edad en el que se concentra el contrato temporal es de 15 a 35 años de edad con un porcentaje en conjunto del 34,40%, mientras que el contrato indefinido se encuentra más en personas de una edad comprendida entre 36 a 65 años de edad que en conjunto representan el 25,64%.

Es claro que al tener un menor nivel de instrucción (Ninguna, Centro de Alfabetización, Educación Básica, Educación Media) las personas se concentran en una relación laboral con contrato temporal, mientras que cuando la persona tiene nivel de instrucción mayor (Superior no universitario, superior, post-grado) se agrupan en contratos indefinidos; tomando en cuenta que en general, del universo analizado el porcentaje de individuos que presentan un nivel de educación superior no universitario y de post-grado tan sólo corresponde al 3,53%. Por otro lado, el contrato temporal se concentra en los individuos solteros, mientras que

los individuos casados tienen mayor tendencia a tener contrato indefinido.

Sexo	Contrato Indefinido		Contrato Temporal		Total	
	Categoría	Número	%	Número	%	Número
Hombre	17447	26,18	23152	34,73	40599	60,91
Mujer	13663	20,50	12393	18,59	26056	39,09
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.1: Distribución por tipo de contrato de la variable *Sexo*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Edad	Contrato Indefinido		Contrato Temporal		Total	
	Categoría	Número	%	Número	%	Número
15-25	5484	8,23	13102	19,66	18586	27,88
26-35	8535	12,80	9822	14,74	18357	27,54
36-45	7565	11,35	6648	9,97	14213	21,32
46-55	6338	9,51	4062	6,09	10400	15,60
56-65	3188	4,78	1911	2,87	5099	7,65
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.2: Distribución por tipo de contrato de la variable *Edad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Nivel de Instrucción	Contrato Indefinido		Contrato Temporal		Total	
	Categoría	Número	%	Número	%	Número
Ninguna	157	0,24	662	0,99	819	1,23
Centro de Alfabetización	44	0,07	127	0,19	171	0,26
Educación Básica	7295	10,94	16611	24,92	23906	35,87
Educación Media	8671	13,01	10486	15,73	19157	28,74
Superior no Universitaria	672	1,01	302	0,45	974	1,46
Superior Universitaria	13142	19,72	7106	10,66	20248	30,38
Post-grado	1129	1,69	251	0,38	1380	2,07
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.3: Distribución por tipo de contrato de la variable *Nivel de Instrucción*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Estado Civil	Contrato Indefinido		Contrato Temporal		Total	
	Categoría	Número	%	Número	%	Número
Casado / Unión Libre	18690	28,04	16785	25,18	35475	53,22
Soltero	8861	13,29	14744	22,12	23605	35,41
Separado /Divorciado	3035	4,55	3533	5,30	6568	9,85
Viudo	524	0,79	483	0,72	1007	1,51
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.4: Distribución por tipo de contrato de la variable *Estado Civil*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Características Laborales

En las tablas A.5, A.6, A.7, A.8 y A.9, se encuentra la información desglosada en lo que respecta a las variables que hacen referencia a las características laborales que discriminan de mejor manera la proporción de contratos temporales vs. contratos indefinidos.

Dentro del análisis se puede evidenciar que tanto el contrato temporal como el indefinido, se concentran en los empleados privados, es decir, la distribución de las individuos en las categorías de ocupación es proporcional tanto para el contrato temporal como para el contrato indefinido.

En lo que respecta a la antigüedad, la capacitación que recibe el empleado en las empresas y el tamaño del establecimiento, se puede destacar que el contrato temporal se concentra cuando el individuo tiene una antigüedad menor igual a un año, se encuentra trabajando en una empresa de menos de 100 trabajadores y no recibe ningún tipo de capacitación.

Ocupación	Contrato Indefinido		Contrato Temporal		Total	
Categoría	Número	%	Número	%	Número	%
Empleado/Obrero Privado	18781	28,18	21578	32,37	40359	60,55
Empleado/Obrero de Gobierno	10490	15,74	2997	4,50	13487	20,23
Jornalero o Peón	145	0,22	7788	11,68	7933	11,90
Empleado(a) Doméstico(a)	1581	2,37	2832	4,25	4413	6,62
Empleado Tercerizado	113	0,17	350	0,53	463	0,69
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.5: Distribución por tipo de contrato de la variable *Ocupación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Antigüedad	Contrato Indefinido		Contrato Temporal		Total	
Categoría	Número	%	Número	%	Número	%
Mayor a un año	26450	39,68	19719	29,58	46169	69,27
Menor igual a un año	4660	6,99	15826	23,74	20486	30,73
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.6: Distribución por tipo de contrato de la variable *Antigüedad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Jornada Laboral	Contrato Indefinido		Contrato Temporal		Total	
Categoría	Número	%	Número	%	Número	%
Jornada Completa	28269	42,41	27632	41,46	55901	83,87
Jornada Parcial	2841	4,26	7913	11,87	10754	16,13
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.7: Distribución por tipo de contrato de la variable *Jornada Laboral*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Tamaño del Establecimiento	Contrato Indefinido		Contrato Temporal		Total	
Categoría	Número	%	Número	%	Número	%
Menos de 100	14583	21,88	28469	42,71	43052	64,59
100 y mas	16527	24,79	7076	10,62	23603	35,41
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.8: Distribución por tipo de contrato de la variable *Tamaño del Establecimiento*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Capacitación	Contrato Indefinido		Contrato Temporal		Total	
Categoría	Número	%	Número	%	Número	%
No	18700	28,05	32238	48,37	50938	76,42
Si	12410	18,62	3307	4,96	15717	23,58
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.9: Distribución por tipo de contrato de la variable *Capacitación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Características Geográficas

En la tabla A.10, se encuentra la información desglosada en lo que respecta a la variable que hace referencia a residencia que discriminan de mejor manera la proporción de contratos temporales vs. contratos indefinidos.

Se puede notar que la región resto de la Costa presenta una mayor agrupación del contrato temporal, mientras que la región resto de la Sierra presenta una distribución similar entre contratos temporales e indefinidos.

Ciudades Autorepresentadas y Regiones	Contrato Indefinido		Contrato Temporal		Total	
	Número	%	Número	%	Número	%
Resto de la Sierra	8449	12,68	8623	12,94	17072	25,61
Resto de la Costa	5307	7,96	10240	15,36	15547	23,32
Guayaquil	3583	5,38	5247	7,87	8830	13,25
Quito	4142	6,21	2715	4,07	6857	10,29
Cuenca	3526	5,29	1531	2,30	5057	7,59
Machala	1606	2,41	3266	4,90	4872	7,31
Ambato	2754	4,13	1711	2,57	4465	6,70
Amazonía	1671	2,51	1948	2,92	3619	5,43
Zonas no delimitadas	72	0,11	264	0,40	336	0,50
Total	31110	46,67	35545	53,33	66655	100,00

Tabla A.10: Distribución por tipo de contrato de la variable *Ciudades Autorepresentadas y Regiones*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

B. Validación Cruzada usando K grupos (K-fold cross validation)

Dado un conjunto m de datos, la técnica de validación cruzada usando K grupos desarrollada en [37] y [39], consiste en:

1. Dividir a los datos en K subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto ($K - 1$) como datos de entrenamiento.
2. Para $i = 1, \dots, K$:
 - Entrenar el clasificador usando todos los datos que no pertenecen al subconjunto i .
 - Probar el clasificador de todos los datos en el subconjunto i .
 - Calcular n_i , el número de datos en el subconjunto i que fueron mal clasificados.
3. Estimar el error de clasificación:

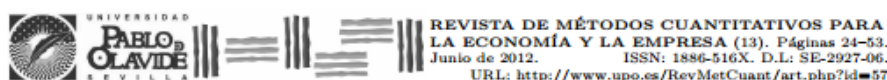
$$E = \frac{\sum_{i=1}^k n_i}{m}$$

El proceso de validación cruzada es repetido durante K iteraciones, con cada uno de los posibles subconjuntos de datos de prueba.

C. Referencia de Publicaciones

Este anexo tiene como objetivo indicar los trabajos de investigación publicados, los cuales fueron realizados durante el periodo en el que se cursó, de manera regular, la maestría y durante la realización del proyecto de titulación.

Publicación (2012)



Determinantes de la temporalidad en el mercado laboral ecuatoriano

CHÁVEZ, YANNIRA

Instituto Nacional de Estadística y Censos, Quito (Ecuador)

Correo electrónico: yannira_chavez@inec.gob.ec

MEDINA, PAÚL

Instituto Gregorio Millán, Universidad Carlos III de Madrid (España)

Correo electrónico: plmedina@espe.edu.ec

RESUMEN

Las diferencias que existen dentro del mercado laboral ecuatoriano, en las diferentes ramas de actividad, evidencian los niveles de formación que los trabajadores deben tener para evitar efectos negativos. Por tal motivo, en este estudio se analiza el papel que las características personales, laborales y geográficas desempeñan en la probabilidad de obtener un contrato temporal frente a uno indefinido. El análisis es realizado por rama de actividad, para determinar qué características posibilitarían la existencia del contrato temporal en cada una de ellas. Para lograr este objetivo, se estiman modelos de regresión logística utilizando los datos de la Encuesta de Empleo, Desempleo y Subempleo desde el 2º trimestre del año 2007 al 2º trimestre del año 2010, elaborada por el Instituto Nacional de Estadística y Censos (INEC).

Palabras clave: contrato temporal; contrato indefinido; regresión logística; rama de actividad.

Clasificación JEL: J23; J24; J41.

MSC2010: 00A71; 03C98; 93E10.

Figura C.1: Publicación Revista de métodos cuantitativos para la economía y la empresa (2012).

Publicación (2013)

Proceedings 59th ISI World Statistics Congress, 25-30 August 2013, Hong Kong (Session IPS017)

p.163

Quantification of losses caused by delinquency in Ecuador

Yannira Chávez^{1,2}, Patricia Cortez^{3,4}, Paúl Medina^{5,6}

^{1,3}National Institute of Statistics and Censuses, Quito, Ecuador

⁵Army Polytechnic School, Quito, Ecuador

²Corresponding author: Yannira Chávez, e-mail: yannira_chavez@inec.gob.ec

⁴Corresponding author: Patricia Cortez, e-mail: patricia_cortez@inec.gob.ec

⁶Corresponding author: Paúl Medina, e-mail: plmedina@espe.edu.ec

Abstract

This paper quantifies the economic cost of several crimes in Ecuador. In particular, we looked at thefts from vehicle, car accessories and homes, robbery of individual, fraud and intimidation (threat and injuries). To do so, we developed a model to estimate expected and unexpected losses, using Loss Distribution Approach (LDA) methodology. We use data from the Survey of Victimization and Perceptions of Insecurity 2011, carried out by the *National Institute of Statistics and Censuses* (INEC, Spanish acronym) and Administrative Records from the *National Police General Direction of Operations*.

Keywords: cost, economic impact, Loss Distribution Approach.

Figura C.2: Publicación en Proceedings 59th ISI World Statistics Congress (2013).