

Bravo, F. y Pinto, C. (2008). Modelos predictivos de la probabilidad de insolvencia en microempresas chilenas. *Contaduría Universidad de Antioquia*, 53, 13-52.

# Modelos predictivos de la probabilidad de insolvencia en microempresas chilenas

***Fernando Bravo Herrera***

Profesor Asociado, Departamento de Administración, Facultad de Economía y Negocios, Universidad de Chile.  
Diagonal Paraguay 257, piso 11, oficina 1106.  
[fbravo@unegocios.cl](mailto:fbravo@unegocios.cl)

***Cristian Pinto Gutiérrez***

Ingeniero Comercial y Magister en Finanzas, Departamento de Administración, Facultad de Economía y Negocios, Universidad de Chile.  
Diagonal Paraguay 257, piso 11, oficina 1106.  
[crpinto@fen.uchile.cl](mailto:crpinto@fen.uchile.cl)

***Clasificación JEL: C45, G21 D82 M1***

---

### ***Modelos predictivos de la probabilidad de insolvencia en microempresas chilenas***

**Resumen:** este trabajo empírico realiza una exploración, revisión y análisis de las principales teorías que modelan un tipo de fracaso empresarial: el no pago de deuda, de tal manera que pueda conformarse un modelo predictivo de impagos, propio de las microempresas chilenas, utilizando una base de 10.519 créditos para 9.649 microempresarios, durante el periodo enero del 2004 a julio del 2006. Se utilizaron dos modelos no-paramétricos del campo de la inteligencia artificial: el método de inducción de reglas rough set y la metodología de redes neuronales artificiales. Ellos fueron comparados con un modelo estadístico tradicional de regresión logística (logit). Para identificar cuál es el modelo que mejor predice el no pago en microempresas, se hizo uso de atributos múltiples que evalúan de manera más precisa el estado económico de este tipo de empresas según sus propias características. Las conclusiones del estudio apuntan a que se pueden lograr capacidades predictivas del orden del 71.9% usando la metodología de redes neuronales.

**Palabras clave:** microempresas, rough set, redes neuronales, fracaso empresarial, asimetrías de información.

### ***Insolvency probability predictive models for Chilean micro enterprises***

**Abstract:** this empirical work offers an exploration, revision, and analysis of the main theories modeling a type of enterprise failure: debt non-payment, so that we manage to have a non-payments predictive model suitable for Chilean micro enterprises using a base of 10,519 credit loans to 9,649 micro enterprises, for the January2004-to-July2006 period. We used non-parametric models from the artificial intelligence field: rough set induction method and artificial neural networks methodology. They were compared with a logistic regression (logit) traditional statistical model. To identify the non-payment best predicting model, we used several attributes that more accurately evaluated the economic state of this type of enterprises according to their own characteristics. Reached conclusions point at the fact that we can have predictive capabilities of about 71.9% using neural networks methodology.

**Keywords:** micro enterprises, rough set, neural networks, enterprise failure, information asymmetries.

### ***Modèles prédictifs de la probabilité d'insolvabilité des micro-entreprises chiliennes***

**Résumé:** ce travail empirique offre une exploration, révision et analyse des principales théories qui modèlent un type d'échec d'entreprise: le non paiement de dette, de sorte que l'on puisse construire un modèle prédictif d'impayés, propre aux micro-entreprises chiliennes, en employant une base de 10,519 crédits pour 9,649 chefs de micro-entreprises pendant la période janvier2004-juillet2006. On a employé deux modèles non paramétriques du champ de l'intelligence artificielle: la méthode d'induction de règles rough set et la méthodologie des réseaux neuronaux artificiels. Ils ont été comparés avec un modèle statistique traditionnel de régression logistique (logit). Afin d'identifier le modèle qui prédit le mieux le non paiement dans les micro-entreprises, on a employé de divers attributs qui évaluent d'une manière plus précise l'état économique de ce type d'entreprises selon leurs propres caractéristiques. Les conclusions de l'étude montrent que l'on peut acquérir des capacités prédictives de l'ordre de 71.9% en employant la méthodologie des réseaux neuronaux.

**Mots clés:** micro-entreprises, rough set, réseaux neuronaux, échec d'entreprise, asymétries d'information.

# Modelos predictivos de la probabilidad de insolvencia en microempresas chilenas

*Fernando Bravo Herrera y Cristian Pinto Gutiérrez*

*Primera versión recibida Septiembre de 2008 – Versión final aceptada Noviembre de 2008*

## I. Introducción

La identificación de los aspectos que condicionan el éxito de las microempresas y la determinación del comportamiento que deben adoptar para mejorar su rendimiento, es una preocupación central de todos aquellos que participan en la formulación de políticas públicas, como de aquellos que participan en la dirección y gestión de empresas privadas que tienen un interés en éstas como principales clientes.

En la última década las empresas chilenas han avanzado aceleradamente hacia su constitución en sujetos de crédito de alto interés para la banca comercial, aumentando su participación absoluta y relativa en el crédito concedido por ésta al conjunto de las empresas chilenas. Así lo sugiere la evidencia empírica. Diversas investigaciones<sup>1</sup> reportan que, en promedio, un 40% de las microempresas tiene acceso al crédito bancario, cifra muy superior al estándar latinoamericano y similar a los niveles de los países desarrollados. Incluso las microempresas reciben más del 9% del crédito total, lo que excede largamente su participación en las ventas (3.4 por ciento) (Chile emprende, SERCOTEC, 2005). Aunque los datos muestran que las microempresas han mejorado su acceso al crédito en los últimos años, existen dificultades con la calidad del acceso a estos servicios, y a los términos y condiciones en que esto se produce. En efecto, como el riesgo de impago es más alto a medida que decrece el tamaño de la empresa, las microempresas tienden a pagar tasas más altas aún en el caso de que el banco tenga la misma información que el deudor. Este mayor riesgo es aumentado por las *asimetrías de información*,

---

<sup>1</sup> Ver, por ejemplo, Cabrera y otros. (2002) y Román (2003).

que generan importantes problemas al operar en contra del acceso fluido de las microempresas al crédito. Esto surge porque en todo momento conviven microempresas rentables con otras débiles que van a fracasar, y son los dueños quienes tienen mejor información que los prestamistas sobre sus probabilidades de sobrevivencia. Esta realidad tiene como consecuencia más grave el racionamiento del mercado. En verdad, las microempresas generan peor información que las empresas medianas o grandes, en parte, porque los requerimientos legales son menos exigentes, pero sobre todo, porque *generar información es un costo fijo que las microempresas no pueden pagar*. La respuesta racional del sistema financiero es operar con aquellas con mayores márgenes, obligar al microempresario a comprometer garantías y/o otorgarle créditos a plazos cortos<sup>2</sup>.

La falta de buena información sobre las microempresas contribuye a profundizar los problemas causados por la asimetría de información, a saber: *la selección adversa y el riesgo moral*.

Esta investigación tiene como objetivo analizar las principales teorías y modelos predictivos de un tipo de fracaso empresarial: el fallo al pagar deuda, de forma que se pueda encontrar el tipo de modelo con un mayor grado de predictividad en las microempresas chilenas. Además, el estudio pretende identificar los aspectos que condicionan la capacidad de la microempresa para efectivamente ser capaz de pagar la deuda. Nos pareció que la manera más acertada de alcanzar estos objetivos era haciendo uso de dos modelos no-paramétricos provenientes del campo de la inteligencia artificial: el método de inducción de reglas *rough set* y la metodología de redes neuronales artificiales.

La comparación de estas dos herramientas con la regresión logística (*logit*) nos permitió conformar el modelo predictivo que clasificara y predigiera con mayor grado de confiabilidad, la posibilidad de quiebra de las microempresas chilenas. Es importante señalar que para identificar cuál era el modelo que mejor predice la incapacidad de la generación de pagos al interior de las microempresas, se hizo uso de atributos múltiples, como ratios financieros, indicadores macroeconómicos y otros indicadores discretos, que señalan el estado económico de este tipo de empresas según sus propias características.

Este estudio empírico fue aplicado a 9.649 microempresas, que obtuvieron crédito de una de las instituciones bancarias chilenas más importantes en micro

---

2 Al respecto, debemos mencionar que en Chile existe crédito a largo plazo para financiar activos que se dan en garantía, por ejemplo, los microempresarios taxistas tienen un acceso bastante expedito al crédito. Sin embargo, no existe crédito contra flujos futuros, porque sería imprudente prestarle a plazos más largos a micro y pequeñas empresas sin más garantías que los flujos futuros esperados. Los plazos más largos atraerían a una proporción muy grande de empresas que fracasarán y eso haría que las que tienen éxito deban pagar tasas más altas para compensar las pérdidas.

créditos, durante el periodo enero del 2004 a julio del 2006. Las conclusiones del estudio apuntan a que se pueden lograr capacidades predictivas del orden del 71.9 por ciento usando la metodología de redes neuronales. Ninguno de los modelos estimados a través de las otras metodologías logra esta eficiencia. Además, se encontró que las variables contables y financieras del negocio no son suficientes por sí solas, y que el mejor predictor de la probabilidad de pago de las empresas de tamaño micro, es la historia del crédito personal de su dueño, las condiciones de financiamiento y sus variables sociodemográficas. En cuanto a la teoría *rough set*, aún cuando el desempeño fue en general inferior al logrado por las otras metodologías, los resultados son satisfactorios al encontrar un 66.2 por ciento de capacidad predictiva promedio. Si consideramos, además, su rapidez para procesar gran cantidad de información, tanto cualitativa como cuantitativa, y su capacidad para generar reglas de decisión fácilmente comprensibles, constituye una alternativa muy competitiva con la técnica más eficiente anterior.

A continuación se presenta el marco teórico de este trabajo, en el que se analizan teorías que explican cómo una empresa cae en cesación de pagos, una caracterización de las microempresas en Chile y una descripción detallada de los modelos de clasificación y predictivos de la probabilidad de impago empresarial que se utilizaron. Luego, se aborda la aplicación de la metodología para la confección de la base de datos, selección de variables y distintas muestras de empresas que se usaron. Finalmente, se presentan y discuten los resultados de cada uno de los modelos aplicados a todas las muestras. Se termina con las conclusiones del estudio.

## II. Teoría y evidencia

Es importante comprender que el fracaso empresarial no necesariamente da como resultado el colapso y la disolución de una empresa. En un sentido económico, el fracaso empresarial puede significar que una empresa está perdiendo dinero, es decir, que el valor presente de los flujos proyectados de efectivo de la empresa es inferior al valor total de su pasivo y patrimonio; o cuando sus proyecciones de rentabilidad no han quedado satisfechas. Aunque, por cierto, el fracaso empresarial puede además referirse más propiamente a una insolvencia.

De acuerdo a Bravo (2006) la insolvencia puede asumir dos formas:

1. Incumplimiento técnico donde una empresa deja de satisfacer una o más condiciones de sus convenios de deuda.
2. Insolvencia técnica que ocurre cuando los flujos de efectivo son insuficientes para satisfacer los pagos de intereses o el reembolso del principal en fechas especificadas. Es decir, no puede cancelar sus deudas al vencimiento.

Vale decir que empíricamente lo observado por los tomadores de decisiones en los bancos es sólo el no pago de las obligaciones de los empresarios, y no así la causa de la insolvencia (es decir que ésta sea un incumplimiento técnico o una insolvencia técnica). Es por ello que cuando nos refiramos a insolvencia durante el desarrollo de este trabajo, nos estamos refiriendo al hecho concreto del no pago, ya sea por incapacidad de flujos o bien por otro tipo de razones desconocidas.

Sin embargo, cuando una empresa se vuelve incapaz de superar sus dificultades y, por ende, ya no puede cumplir con sus obligaciones existen varias alternativas de reorganización financiera. Pero para que sean factibles algunas de las alternativas de reorganización financiera, el deudor debe mostrar capacidad para recuperarse y las condiciones generales de negocios deben ser favorables para la recuperación.

En un entorno que está cambiando continuamente el desempeño de las grandes organizaciones va sistemáticamente perdiendo eficiencia hasta que resultan reemplazadas, y por tanto, empresas inviables y de baja productividad son desplazadas por nuevas firmas permitiendo la evolución industrial. Este modelo de dinámica industrial se remonta a la década de los sesenta, cuando Solow (1960) formula su modelo de “*vintaje*”, el cual supone que parte fundamental del progreso tecnológico se encuentra incorporado en las nuevas empresas dotadas de bienes de capital de última generación.

No obstante, un enfoque más apropiado en nuestro caso, es la visión opuesta, establecida por lo que se conoce como *modelos de selección*: el grueso de los nuevos entrantes no serán capaces de sobrevivir siquiera al muy corto plazo, ya que ellos entran al mercado basados en sus expectativas que una innovación o nuevas ideas los harían viables en el proceso de competencia. Sin embargo, al cabo de unos pocos periodos productivos se dan cuenta que estas expectativas estaban sobredimensionadas, que el experimento no es rentable y saldrán del mercado; por ejemplo Jovanovic (1982) provee una evidencia al respecto: empresas pequeñas tienen mayor probabilidad de quiebra que empresas grandes. Posteriormente, Hopenhayn (1992) desarrolla y analiza un modelo que muestra que las tasas de sobrevivencia serán mayores para las firmas más antiguas.

El estudio empírico de Crespi (2003), muestra que en Chile las empresas que crecen y sobreviven son en general las de los estratos de mayor tamaño. Son muy pocas las microempresas que se mueven en forma ascendente sobre el tiempo sugiriendo que se está frente a una dinámica de *puerta giratoria* de crecimiento.

### *II.1. Revisión de la literatura sobre la predicción de impagos.*

Este proceso de selección está determinado por una serie de factores agregados relacionados con la situación macroeconómica general, las condiciones de financiamiento y el comportamiento de las empresas. Desde este punto de vista, son varios los factores específicos que conducen a la

generación de impagos en una empresa, y develarlos ha sido la motivación de los estudiosos del campo.

El primer acercamiento científico del proceso de quiebra e impagos conduce a los investigadores a utilizar la información contable y modelos matemáticos y estadísticos. Entre los estudios más destacados se encuentran los trabajos pioneros de Beaver (1966) basado en el enfoque de análisis discriminante univariado y el modelo multivariable de Altman (1968), el que se convirtió en el modelo dominante predictor de insolvencias y estándar de comparación para otros modelos posteriores. En las décadas de los ochentas y noventas, la mayoría de los autores utilizaron modelos más complejos, pero al mismo tiempo más precisos, para determinar la probabilidad de quiebra de empresas; especialmente los modelos *Logit* (Ohlson, 1980; Platt y Platt, 1990; Platt y otros, 1994), el análisis *Probit* (Basch y Montenegro, 1989) y, en la última década, las Redes Neuronales (Wilson y Sharda, 1995; Odom y Sharda, 1990; Rahimian y otros, 1993; Atiya, 2001). El enfoque común de estos modelos es la revisión de la literatura identificando un gran conjunto de variables financieras predictivas para luego desarrollar un reducido conjunto de variables predictivas. Las razones contables de liquidez, actividad, rentabilidad y endeudamiento fueron los indicadores más importantes de estos modelos.

Más recientemente, una avanzada técnica no-paramétrica, la teoría *rough set*, la cual tiene sus fundamentos en la teoría matemática de conjuntos, ha sido aplicada a la solución de estos problemas. La teoría *rough set* fue desarrollada por Pawlaw (1991), no requiere que los datos satisfagan ningún tipo de requisito o hipótesis, no se ve afectada por la presencia de observaciones atípicas y el modelo final obtenido consiste en un conjunto de reglas de decisión que las hacen fácilmente comprensibles. Por lo anterior, es una buena alternativa para el estudio de la insolvencia cuando se carece de información contable fiable o donde simplemente no existe. La aplicación del enfoque *rough set* en la predicción de modelos de este tipo fue investigada por Dimitras y otros, (1999) en una muestra de firmas griegas. Posteriormente, otros estudios en mercados distintos han aplicado la metodología *rough set* a la predicción de la incapacidad de pago; Mckee (2000) en Estados Unidos y Beynon y Peel (2001) en Gran Bretaña. Dichas investigaciones no hicieron más que confirmar que los modelos de predicción basados en el enfoque *rough set* constituyen una importante alternativa a los modelos estadísticos clásicos en la predicción de los impagos.

### **III. Caracterización de la microempresas chilenas**

#### *III.1. Generalidades*

La masa empresarial chilena incorpora una amplia gama de pequeñas, medianas y grandes empresas, acompañadas por una aún mayor masa de

microempresas y trabajadores por cuenta propia. Al año 2004 había 707 mil 634 empresas formales de carácter micro, pequeño, mediano o grande. Del total de unidades económicas formales las microempresas constituyen el 81%, porcentaje que se mantiene con poca variación durante el periodo 1999 a 2004<sup>3</sup> (Ver Tabla 1).

Tabla 1. Distribución por tamaño de empresas formales

Año	Micro	%	Pequeña	%	Mediana	%	Grande	%	Total
1999	527.481	80.9%	103.891	15.9%	14.198	2.2%	6.344	1.0%	651.914
2000	552.125	84.0%	85.873	13.1%	13.098	2.0%	5.966	0.9%	657.062
2001	567.861	84.2%	87.186	12.9%	13.39	2.0%	6.187	0.9%	674.624
2002	569.836	83.8%	89.872	13.2%	13.833	2.0%	6.546	1.0%	680.087
2003	570.544	81.8%	105.624	15.1%	14.577	2.1%	6.868	1.0%	697.613
2004	571.535	80.8%	112.731	15.9%	15.748	2.2%	7.620	1.1%	707.634

Fuente: datos extraídos del libro "La situación de la pequeña y microempresa en Chile" y basados en un estudio de FUNDES.

Las microempresas formales muestran una participación en las ventas totales, entre el periodo 1999 y 2003, de alrededor de 3,4 % (Ver Tabla 2).

Tabla 2. Ventas anuales por tamaño de empresas formales (miles de millones año 2003)

Año	Micro	%	Pequeña	%	Mediana	%	Grande	%	Total
1999	4.386	4.2%	12.457	12.0%	11.142	10.7%	76.087	73.1%	104.072
2000	3.407	3.4%	10.534	10.5%	10.263	10.3%	75.667	75.8%	99.871
2001	3.486	3.1%	10.78	9.5%	10.581	9.3%	88.594	78.1%	113.441
2002	3.541	2.9%	11.056	9.0%	10.944	8.9%	96.914	79.1%	122.455
2003	4.557	3.4%	12.793	9.5%	11.766	8.8%	105.323	78.3%	134.439

Fuente: datos extraídos del libro "La situación de la pequeña y microempresa en Chile" y basados en un estudio de FUNDES.

Si bien las microempresas abarcan un conjunto de negocios con alto grado de heterogeneidad, es posible describirlas en función de ciertas características comunes:

- Generalmente corresponden a negocios de carácter familiar o trabajadores auto empleados (Focus, 2005).
- Son organizaciones productivas de bajos rendimientos, infraestructura insuficiente e inadecuada, gestión conservadora y bajos salarios.
- Generalmente ineficientes en términos de abastecimiento de materias primas, comercialización y manejo contable y financiero.
- El empleo generado por la microempresa es más bien precario, sin adecuada protección social, bajas remuneraciones y alta rotación laboral.

<sup>3</sup> A este porcentaje habría que agregar el aporte que hacen las microempresas informales, cuya magnitud estimada alcanza las 748 mil unidades productivas (año 2004).



- Como promedio la venta de las microempresas<sup>4</sup> chilenas es de 456 UF anuales, o sea, US\$16.000 aprox. (Roman, 2003), monto muy pequeño que genera inestabilidad a sus empleados y propietarios.
- Se caracterizan por su escaso peso de mercado.

Sin perjuicio de lo anterior, debido al grado de heterogeneidad que existe entre las microempresas, es posible diferenciarlas en relación a los potenciales de crecimiento y el efecto de los créditos entre las microempresas. Así, tenemos empresas de baja potencialidad, como en el sector comercio y de alta potencialidad, como es el sector industrial. Aunque la evidencia indica que las microempresas están presentes en todas las actividades económicas, su mayor incidencia está en aquellos estratos de menor potencialidad y barreras de entrada, como es el caso del comercio. Las cifras indican que el 77% de las microempresas se concentra en 4 sectores (comercio, servicios, transporte y agricultura)<sup>5</sup>.

### III.2. Problema de financiamiento bancario de las microempresas

Las características anteriores de las microempresas tienen un impacto considerable sobre el acceso de éstas al financiamiento y sobre las relaciones que establecen con el sector financiero. De particular importancia, el riesgo de cartera de este estrato es efectivamente superior al de las empresas de mayor tamaño, como se aprecia en la Tabla 3.

El hecho que una institución financiera esté dispuesta a prestar dinero a una firma conlleva información valiosa para el resto e implica una externalidad positiva. Pero un segundo prestamista (beneficiado por la externalidad anterior), puede imponer una externalidad negativa sobre el primero, puesto que la probabilidad de “default” es función del monto total adeudado (altos intereses o garantías). Esta situación tiende a generar un sesgo en contra de los préstamos a empresas nuevas, o sin historia, como en el caso de las microempresas, que se caracterizan, por una alta rotación y corto ciclo de vida.

Tabla 3. Evolución de empresas formales con deuda vencida respecto a empresas endeudadas

Año	Micro	Pequeña	Mediana	Grande	Total
1999	9.0%	8.8%	6.1%	3.5%	8.7%
2000	9.0%	8.1%	5.3%	2.6%	8.6%
2001	8.3%	7.7%	5.4%	2.9%	8.0%
2002	9.0%	8.4%	5.5%	3.1%	8.7%
2003	8.9%	9.0%	5.3%	2.2%	8.7%

Fuente: datos extraídos del libro “La situación de la pequeña y microempresa en Chile” de acuerdo a datos de la SBIF y el SII.

- 4 Las “Microempresas” fueron definidas como aquellas unidades económicas formales con menos de 2.400 UF de ventas netas anuales, o sea, unos US\$ 85.000 aproximadamente, según clasificación CORFO.
- 5 Las cifras referidas corresponden a un estudio de FUNDES, realizado por encargo de SERCOTEC el año 2005.

Por otra parte, los costos fijos de transacción limitan el acceso al crédito de las microempresas a los mercados financieros. Así una cartera compuesta por numerosos créditos de menor cuantía, tiene, proporcionalmente, un mayor costo relativo por este concepto. Esto unido a la carencia de información estandarizada y confiable, y de proyectos bien informados y en condiciones de ser evaluados, por parte de las microempresas, generan mayores costos de estudio de cada proyecto de créditos. Como consecuencia directa del mayor riesgo de la cartera y de los elevados gastos de apoyo, el sistema financiero exige a las microempresas mayores *spreads* (intereses) respecto de empresas de mayor tamaño.

#### IV. Modelos de clasificación y predictivos de los impagos

##### IV.1. La teoría *rough set*

La *teoría rough set* es una poderosa herramienta matemática para manejar la imprecisión y la incertidumbre inherente al proceso de toma de decisiones. Fue originalmente desarrollada en la década de los ochenta por Pawlak (1982, 1984, 1991) y desde entonces esta teoría ha sido bien estudiada por numerosos investigadores que han introducido importantes innovaciones al enfoque clásico<sup>6</sup>. El concepto de *rough set* se fundamenta en el supuesto de que a cada objeto en la naturaleza que estemos considerando, podemos asociarle algún conocimiento o información. *Objetos caracterizados por la misma información son indiscernibles (similares) con respecto a la información disponible*, y son tratados como idénticos o similares. Por ello, la teoría *rough set* es útil cuando las clases en las que han de clasificarse los objetos son imprecisas, sin embargo pueden aproximarse mediante conjuntos precisos (Nurmi y otros, 1996).

Pawlak explica:

[...] una de las principales ventajas de la teoría *rough set* es que ésta no necesita información preliminar o adicional de ningún tipo sobre los datos, tales como distribución de probabilidad en estadísticas o grado o probabilidad de pertenencia en la teoría de conjuntos difusos (*fuzzy set theory*). (Pawlak, 1991, p. 1)

Esta sección presenta los principales conceptos de la teoría *rough set* siguiendo las referencias de Pawlak (1991, 2000); Dimitras y otros, (1999) y Komorowski y otros, (1999).

*Sistema de información.* La teoría *rough set* asume la representación del conocimiento sobre los objetos en forma de un sistema de información.

---

6 Aunque existen en la actualidad extensiones de dicha teoría en esta investigación sólo presentamos el enfoque clásico. Para una detallada revisión de la teoría *rough set*, sus aplicaciones y extensiones sugerimos referirse a Komorowski y otros; (1999), Greco y otros; (1998, 2001) y Ziarko (1993).

Tabla 4. Ejemplo de una tabla de decisión

$U$	Atributos de condición						Atributo de decisión
	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$d$
$x_1$	1	0	1	1	0	1	L
$x_2$	1	0	0	1	0	0	L
$x_3$	0	0	1	0	0	0	L
$x_4$	1	0	1	1	0	1	H
$x_5$	0	0	0	0	1	1	H
$x_6$	1	0	1	1	0	1	H
$x_7$	0	0	0	0	1	0	H

Fuente: elaboración propia

Un sistema de información es ilustrado en la Tabla 4, en las filas de la tabla se indican el conjunto de objetos  $U = \{x_1, x_2, \dots, x_7\}$ . Las columnas denotan los atributos  $A = \{a_1, a_2, \dots, a_7\}$  de estos objetos y la variable de decisión  $d$ . Las entradas en la tabla son los valores de los atributos o descriptores. Cada fila de la tabla contiene descriptores que representan información correspondiente a un objeto. La relación de no diferenciación ocurrirá si dos objetos para todos los atributos tomasen los descriptores el mismo valor.

*Lower and upper approximation.* Debido a la imprecisión que existe en los datos en el mundo real, siempre existirá conflicto en los datos contenidos en una tabla de decisión. Aquí el conflicto se refiere a dos o más objetos idénticos usando cualquier conjunto de atributos de condición, sin embargo, ellos tienen una diferente clase de decisión. Tales objetos son llamados *inconsistentes*. Esta tabla de información es llamada *tabla de decisión inconsistente*.

En la teoría *rough set*, la *aproximación de conjuntos* es introducida para tratar con esta inconsistencia. Aquellos elementos que no son inconsistentes, se les conoce como una *B-indiscernibility relation* y puede ser denotada por  $IND(B)$ .

$IND(B)$  divide el universo  $U$  en familias de clases de equivalencia  $\{x_1, x_2, \dots, x_7\}$ , las cuales son comúnmente conocidas como *clasificaciones de U* y son denotadas por  $U/IND$ . Una *clase equivalente*  $x_i$  de la relación  $IND(B)$  se denomina *conjunto elemental de B* (*B-elementary set*) y es denotado por  $Y$  si es que este contiene algún elemento  $x$ .

Aunque algunos objetos en una tabla de información no pueden ser exactamente distinguidos dado un conjunto de atributos, estos pueden ser aproximadamente diferenciados. La idea consiste en reemplazar o representar el *rough set* por un par de conjuntos precisos, llamados *aproximación por debajo y por encima* (*B-lower and B-upper approximations*). La aproximación por debajo y por encima de  $x$  puede ser definida, respectivamente, como sigue:

$$\underline{B}X = \bigcup \{Y \in U \setminus IND(B): Y \subseteq X\}$$

$$\overline{B}X = \bigcup \{Y \in U \setminus IND(B): Y \cap X \neq \emptyset\}$$

La aproximación por debajo de  $X$  es la unión de todos los conjuntos elementales que están incluidos en  $X$ , mientras que la aproximación por encima de  $X$  es la unión de todos los conjuntos elementales que tienen una intersección no vacía con  $X$ . En otras palabras, la aproximación por debajo consiste en todos los objetos que seguramente pertenecen al conjunto y la aproximación por encima contiene los objetos que posiblemente pertenecen al conjunto. El conjunto  $BN_B(X) = \underline{B}X - \overline{B}X$  es llamado *frontera de  $X$  ( $B$ -boundary región)* y consiste en el conjunto de elementos que no puede ser con certeza clasificados utilizando el conjunto de atributos. Si la región de frontera de  $X$  es un conjunto vacío, es decir,  $BN_B(X) = \emptyset$ , entonces  $X$  será *crisp (exacto)* con respecto a  $B$ , en caso contrario, es decir, si  $BN_B(X) \neq \emptyset$ , nos referiremos a  $X$  como *rough (inexacto)* con respecto a  $B$ .

*Calidad de clasificación.* La inexactitud de un conjunto es debida a la existencia de una región de frontera. Mientras mayor sea la región de frontera, menor será la precisión del conjunto. Luego, dada la aproximación de una clasificación, al cociente entre el número cardinal de la aproximación por debajo y el de la aproximación por encima se le denomina *precisión*. Este ratio expresa el porcentaje de posibles clasificaciones correctas cuando los objetos son clasificados usando el conocimiento disponible. Si lo que tenemos es una clasificación en varios grupos, para cada clase se pueden calcular las aproximaciones por encima y por debajo. Al cociente de la suma de aproximaciones por debajo de todas las clases dividido por el total de elementos que disponemos se le denomina *calidad de clasificación* de  $A$  por  $B$ . Esta es una medida para describir la inexactitud de la clasificación por aproximación y expresa el porcentaje de objeto que pueden ser correctamente clasificados en las clases de  $A$  empleando los atributos de  $B$ .

$$\gamma_B(A) = \frac{\sum \text{card}(\underline{B}X)}{\text{card}(U)}$$

Si  $\gamma_B(A) = 1$ , entonces la tabla de decisión es consistente. De otra forma es inconsistente.

*Reductos y núcleo.* Al eliminar aquellos atributos que son redundantes, podemos obtener un conjunto reducido de atributos que mantiene la misma *calidad de aproximación* del conjunto original, a lo que le llamamos *reducto*. En una tabla de información puede haber más de un reducto y su intersección nos da el denominado *núcleo de  $A$  ( $B$ -core)*. El núcleo es la parte esencial de  $A$ , la cual no

puede ser eliminada sin afectar la calidad de clasificación de la aproximación de los objetos al interior de las clases elementales de  $B$ .

*Reglas de decisión.* Un sistema de información reducido permite la obtención de reglas de decisión. De hecho ésta es la cuestión más importante del enfoque *rough set*.

Los reductos de los conjuntos de atributos de condición mantienen una importante relación con las clases de decisión. Una tabla de decisión especifica las decisiones (acciones) que deben ser tomadas cuando algunas condiciones son satisfechas. Debido a esta dependencia funcional entre atributos de condición y de decisión, una tabla de decisión puede ser vista también como un conjunto de reglas de decisión. Entonces, una tabla de información reducida puede proveernos reglas de decisión de la forma “*si, ... , entonces*”. La sintaxis de la regla es como sigue:

Si  $f(x, a_1)$  y  $f(x, a_2)$  y ...  $f(x, a_k)$ , entonces  $x$  pertenece a  $DA_1$  o  $DA_2$  o  $DA_k$ , donde  $\{a_1, a_2, \dots, a_k\} \subseteq CA$  y  $DA_1$  o  $DA_2$  o  $DA_k$  son clases de decisión. Si  $n=1$ , entonces la regla es exacta (determinística); de otra forma, ésta es aproximada o ambigua (Greco y otros, 2001), que sería cuando las condiciones pueden conducir a varias posibles decisiones.

Una regla de decisión en  $S$  es expresada como  $\emptyset \rightarrow \psi$ , donde  $\emptyset$  y  $\psi$  son las condiciones y decisiones de las reglas de decisión, respectivamente. Cada regla de decisión  $\emptyset \rightarrow \psi$  se caracteriza por su fuerza (*strenght*), esto es, el número de objetos que satisfacen la parte de la condición de la regla (cubiertos por la regla) y que pertenecen a la clase de decisión sugerida por ella.

$$\sigma_s(\phi, \psi) = \frac{\sup[p_s(\phi, \psi)]}{\text{card}(U)}$$

El conjunto de reglas para todas las clases de decisión se denomina algoritmo de decisión. Así, el modelo final obtenido consiste en un conjunto de reglas de decisión que las hacen fácilmente comprensibles por el analista y permiten justificar y explicar las conclusiones derivadas del análisis de los datos. Además, da la posibilidad al analista de controlar dicho análisis de manera simple. Tal posibilidad no es normalmente ofrecida por las técnicas tradicionales de análisis de datos, añadiendo además una baja carga computacional y posibilidad de interacción con un usuario en modo de diálogo.

#### IV.2. Las Redes Neuronales Artificiales

Los modelos de Redes Neuronales Artificiales corresponden a una elaboración matemática inspirada en el sistema nervioso biológico. Trata de reproducir algunos mecanismos que tiene el cerebro humano para procesar información, reconocer patrones, hacer predicciones o tomar una decisión con base en la experiencia pasada. La literatura destaca que éstas pueden

ser aproximadores de funciones universales aún para funciones no lineales (Hornik y otros, 1989). Lo anterior significa que ellas pueden aproximar automáticamente cualquier forma funcional (lineal o no lineal) que mejor caracterice los datos, permitiéndole a la red extraer más señales a partir de formas funcionales subyacentes complejas. (Hill y otros, 1994)

Durante las últimas décadas se han usado las redes neuronales para la clasificación de agentes económicos tales como compañías de seguros, tenedores de tarjetas de crédito, cierre de cuentas corrientes y, en la situación que acá se aborda, para obtener una estimación de la probabilidad de quiebra de compañías. Entre los primeros autores que han escrito sobre probabilidad de quiebras corporativas usando redes neuronales se encuentran Odom y Sharda (1990); Wilson y Sharda (1994) y Rahimian y otros, (1993). En los últimos años Atiya (2001).

Básicamente, las redes neuronales pueden entenderse como modelos multicuacionales y consisten en varias capas de neuronas en las que los *output* de unos constituyen los *input* de otros. Los componentes de una red neuronal son: (a) *unidades de entrada*, (b) *unidades de salida*, las cuales corresponden al resultado de la red neuronal que puede hacerse a través de una sola neurona o de manera lineal ponderando y sumando las salidas de la última capa; y (c) *unidades ocultas*, que son aquellas en donde se lleva a cabo el procesamiento de los datos. Luego, el problema se centra en determinar los parámetros adecuados de las redes neuronales que contengan la combinación correcta de elementos de procesamiento, tasas de aprendizaje, arquitectura y número de capas con tiempos de entrenamientos aceptables y que posean un buen rendimiento (Parisi y otros, 2003). La red típica suele tener una salida y una o dos capas ocultas con un número cercano de neuronas al número de variables de entrada.

## V. Metodología, datos y selección de variables

El objetivo del presente trabajo es desarrollar un modelo para identificar a las microempresas que tienen una alta probabilidad de fracasar, identificando las variables relevantes a incluir en el modelo de predicción y evaluar la efectividad de los distintos modelos planteados. Vale destacar que este objetivo no incluye el estimar la magnitud de los efectos de las diversas variables en dichas probabilidades, ya que depende en gran medida de la forma de confección y definición de los datos, y de la forma de utilización de los métodos planteados. Por ello, con base en los modelos presentados teóricamente en las secciones anteriores, nuestro objetivo es encontrar aquel modelo o configuración de variables y modelo que permita tener mejor capacidad predictiva.

Una microempresa que no logra acceder al crédito es aquella que no logra demostrar que será capaz de generar los flujos de caja necesarios para cumplir con sus obligaciones financieras, presenta incumplimiento de compromiso

con otros actores del sistema financiero o servicio de impuestos internos, son empresas extremadamente sensibles a fluctuaciones cíclicas de la economía, o bien se trata de empresas informales de bajo potencial. Para ello nos basamos en la información histórica disponible en la base de datos de una de las instituciones bancarias con mayor experiencia en el campo de la microempresa en Chile, construyendo una muestra de corte transversal de microempresas. Es posible además, con el objetivo de asegurar la representatividad de la muestra, demostrar las similitudes entre las microempresas de nuestro estudio con algunos datos existentes sobre el total de microempresas. Sólo para ilustrar, Román (2003) calcula que como promedio las ventas de las microempresas son de 450 UF anuales. El promedio de ventas de nuestra muestra de microempresas es de 400 UF. El mismo Román (2003) asevera que: o las microempresas son de un tamaño muy pequeño, o pagan salarios muy bajos, o bien exhiben una combinación de ambos efectos. Para nuestra muestra, en efecto, es tan bajo el nivel de las ventas de las microempresas y sus ingresos de otras fuentes, que si el total de dichos ingresos por ventas se usasen para pagar las cuotas mensuales del crédito y otros gastos familiares se gastarían en promedio un 37% de sus ingresos por ventas mensuales; si el resto lo dedicaran a salarios, equivaldrían sólo a una capacidad de pago de aproximadamente 3 salarios mínimos mensuales<sup>7</sup>. Eso explica porque más del 60% de las microempresas de nuestro estudio son unipersonales.

La información solicitada correspondió a 4 tipos de datos: características del microempresario, características de la microempresa, atributos del crédito y comportamiento de pago del microempresario. Datos externos correspondientes a variables relacionadas al ciclo económico y condiciones macroeconómicas generales de la economía fueron tomadas de cifras oficiales del Banco Central de Chile. Estas corresponden a la tasas de crecimiento del producto (medido por el IMACEC), tasa de desempleo y tasa de interés (tasa de política monetaria) las cuales fueron asociadas de acuerdo al mes inmediatamente anterior al que la empresa cae en la categoría de impagos<sup>8</sup>.

De la etapa de preproceso de los datos se obtuvo un total de 10.519 créditos para 9.649 microempresarios. Para propósitos de la investigación se mantuvo sólo el último crédito de cada microempresario, por lo que la muestra final quedó conformada por 9.649 microempresas. Como además no se sigue la evolución de la capacidad de pago del microempresario, sino sólo su situación final, esta base está constituida por datos de corte transversal, y no representa datos de panel. Por otra parte, de un total de 192 atributos disponibles de estos

---

7 El salario mínimo vigente a diciembre del 2006 equivale a \$135.000

8 En el caso de la empresa sana, cuando su morosidad fue máxima. Para las que no son morosas se considera el mes anterior a la última cuota de su crédito.

contratos, se llegó a un total de 82 variables en una tabla que fue utilizada en el proceso de transformación de datos.

### *V.1. Identificación de la muestra de entrenamiento, validación y prueba*

Aunque en la práctica financiera de las microempresas, la quiebra no sigue estrictamente un proceso formal, para la definición de empresa con impagos se consideraron dos enunciaciones alternativas. Primero se definió como microempresa con impagos aquella cuyo último crédito se encuentra en cartera castigada y/o se pagó mediante el embargo de los bienes del microempresario. Para la segunda definición, se consideró un indicador de “*insolvencia severa*” medido como mora promedio ponderada superior a 90 días. Cualquiera de estos casos corresponde a una empresa que o bien no cumple con sus obligaciones financieras, o bien adolece de serias dificultades para cumplirlas, representadas en el retraso promedio mayor a 90 días en el pago.

Así se clasificaron 980 microempresas en la categoría insolvente, que corresponde al 10.15% de la muestra final. La muestra de entrenamiento y validación constó de 1404 microempresas, 702 parejas de microempresas de los grupos solvente e insolvente, respectivamente. Para el proceso de validación el conjunto total de objetos de la muestra de entrenamiento es dividido en 2 subconjuntos. Las fases de aprendizaje y clasificación son realizadas con el 70% de la muestra, dejando un subconjunto del 30% como conjunto de validación. La tercera muestra, la de prueba, consiste en 278 empresas del grupo fracasadas y 2.033 del grupo de no fracasadas.

En consideración a que se contó con datos de créditos otorgados entre enero del 2004 a julio del 2006, se construyó el conjunto de entrenamiento y validación con los créditos colocados durante el año 2004. El conjunto de prueba estuvo conformado por las colocaciones de inicio del 2005 hasta julio del 2006.

### *V.2. Selección de variables explicativas*

Las variables explicativas que fueron desarrolladas en nuestros modelos se organizaron en nueve grupos de características<sup>9</sup>:

1. *Variables sociodemográficas*
2. *Condiciones de Financiamiento*
3. *Ingresos y gastos mensuales del microempresario*
4. *Historia y comportamiento de pago*
5. *Consolidación Familiar*
6. *Comportamiento de ahorro y consolidación patrimonial*

---

<sup>9</sup> Estas variables fueron obtenidas luego de un detallado análisis estadístico univariado, donde se seleccionaron las más relevantes objeto de modelación.



7. *Experiencia como microempresario*
8. *Ratios financieros*<sup>10</sup>
9. *Entorno macroeconómico.*

Como se aprecia, además de información financiera de la microempresa se incorporó información sobre el entorno macroeconómico, para controlar el estado de ciclo económico; y sobre el dueño, puesto que en estos casos la decisión de crédito se asemeja a la de un crédito de consumo.

El *análisis discriminante* nos permitió determinar cuáles variables de pronóstico contribuyen más a la diferencia entre los grupos de fracasadas o sanas en función del valor que tomen las variables observadas.

Los resultados obtenidos por el análisis discriminante mostró a las 35 variables más representativas, de las cuales, eliminamos aquellas que no difieren estadísticamente entre los grupos de fracasadas y no fracasadas (5 variables cuyos test *F* no son significativos). En el caso de la metodología *rough set* no fue necesario una nueva selección, ya que por sus características no tendremos problemas de correlación y nos permitirá reducir aun más las variables redundantes. La situación es análoga para las redes neuronales, puesto que tampoco constituye un modelo estadístico. Sin embargo, para la utilización de la técnica *logit*, se hizo necesario el uso del *análisis factorial*, en lugar del análisis discriminante, para garantizar la especificación estadística de los datos.

El *análisis factorial* tiene como objetivo simplificar las múltiples y complejas relaciones que pueden existir entre un conjunto de variables observadas. Examina las correlaciones entre el conjunto de variables o atributos que caracterizan a las microempresas y resume esas variables en un número menor de factores o dimensiones subyacentes<sup>11</sup>. A través de este procedimiento identificamos un nuevo conjunto más pequeño de variables no correlacionadas, que reemplaza el conjunto original de variables continuas correlacionadas, para la inclusión posterior en el análisis *logit*.

## VI. Resultados

### VI.1. Resultados de la aplicación del método *rough set*

El primer resultado obtenido del análisis *rough set* de la tabla de información codificada fue la precisión y la calidad de clasificación, ambas igual a uno. Dado el gran número de variables, las microempresas están muy bien discriminadas entre ellas. Es decir, no hay casos frontera ni hay, por tanto, información imprecisa.

---

10 Para detalles sobre la creación de estas variables ver Anexo 2. Creación de Ratios Financieros

11 Las variables deben ser de nivel intervalo o de razón, no siendo aconsejable el procedimiento para las variables categóricas (tales como región, nivel de educación, etc.).

Luego se construyó el mínimo subconjunto de atributos independientes que asegurasen la misma calidad de clasificación que la totalidad del conjunto, esto es los reductos. Hemos obtenido 11.684 reductos de la tabla de información codificada, cada uno de los cuales contiene entre 14 y 21 atributos, lo que, respecto a los 30 originales, supone una reducción importante, porque nos permitió eliminar, al menos 9 atributos redundantes sin ninguna consecuencia. Las variables que aparecen con más frecuencia en los reductos (aparecen en más del 50%) son: *Tasa de política monetaria (TPM)*, *sexo*, *región*, *estado civil*, *edad*, *actividades sociales*, *formalidad de la microempresa*, *antigüedad en el sector*, *periodicidad del ahorro*, *capital original* y *capacidad de pago* y los ratios financieros *capital de trabajo sobre activos totales (KT/AT)*, *retorno sobre los activos (ROA)*, *cobertura de intereses*, y *rotación de activos* lo que indica que son variables muy discriminatorias en la muestra entre microempresas solventes e insolventes. El núcleo está compuesto de un atributo, *TPM*, esto significa que éste es el único atributo absolutamente necesario para la aproximación de las dos clases de decisión. Esto confirma nuestra hipótesis que señala que *las variables relacionadas con el entorno de la microempresa afectan su probabilidad de sobrevivencia*<sup>12</sup>.

Tabla 5. Atributos en los reductos seleccionados

Reducto	Atributos
Reducto 1	{Región, Sexo, Estado Civil, Edad, Renegociación, Capacidad de pago, Ln(ventas), Cobertura de i, ROA, Antigüedad sector, TPM, Formalidad, Residencia, Sercotec}
Reducto 2	{Región, Sexo, Estado Civil, Edad, Capacidad de pago, Cobertura de i, ROA, Actividades sociales, vehículo, Antigüedad sector, TPM, Formalidad, Residencia, Sercotec}
Reducto 3	{Capital, Región, Sexo, Renegociación, Experiencia potencial, Capacidad de pago, Ln (ventas), Cobertura de i, ROA, KT/AT, Casa propia, Antigüedad sector, TPM, Formalidad}
Reducto 4	{Capital, Región, Sexo, Estado Civil, Edad, Capacidad de pago, Cobertura de i, ROA, Actividades sociales, Antigüedad sector, TPM, Formalidad, Residencia, Sercotec}
Reducto 5	{Región, Sexo, Estado civil, Edad, Capacidad de pago, Cobertura de i, ROA, Razón corriente, Actividades sociales, Periodicidad ahorro, Antigüedad sector, TPM,,Residencia }

Fuente: elaboración propia

12 Existe también una serie de estudios que han mostrado que las magnitudes de muerte de empresas esta influenciada, aunque con cierto rezago, por el estado del ciclo económico (Yamawaky H., 1991; Highfield y Smiley, 1987). Para controlar estas influencias se incluyeron entre las variables, valores rezagados del crecimiento del índice mensual de actividad económica (IMACEC), de la tasa nacional de desempleo y tasa de interés (tasa de política monetaria). Sin embargo, sólo la evolución de la tasa de interés de la economía resultó ser una variable relevante para discriminar entre las empresas que mueren y las que sobreviven, de acuerdo a los resultados del análisis discriminante.

Tabla 6. Lista de frecuencia de los atributos en reductos

Atributo	Frecuencia	%
TPM	11.684	100
Sexo	10.442	89.4
Región	9.993	85.5
Estado civil	9.871	84.5
Edad	8.076	69.1
Actividades...	9.609	82.2
Formalidad...	9.355	80.1
Antigüedad...	10.8	92.4
Periodicidad...	6.69	57.3
Capital...	6.012	51.5
Capacidad...	7.049	60.3
KT/AT	8.676	74.3
ROA	8.594	73.5
Cobertura...	8.468	72.5
Rotación....	7.14	61.1

Fuente: elaboración propia

El gran número de reductos obtenidos implicó un detallado análisis de éstos para obtener aquél que generara un menor número de reglas y que mostrara un buen desempeño predictivo. Los reductos fueron seleccionados teniendo en cuenta los siguientes criterios: a) El reducto debía contener el menor número de atributos que fuese posible (del total de reductos posibles 127 cumplen con este criterio). b) Debía contener los atributos considerados a nuestro juicio más significativos para la evaluación de las microempresas. c) No debía contener variables que estuviesen correlacionadas. Luego, los cinco reductos que hemos elegido se muestran en la Tabla 5.

Elegidos los reductos, el resto de atributos han sido eliminados de la tabla de información codificada y hemos pasado de un sistema inicial de información codificado de 30 columnas, a cinco sistemas de sólo 14. La muestra de 1404 firmas empleadas en esta parte del estudio para derivar el algoritmo de clasificación, puede considerarse como una muestra de aprendizaje utilizada para revelar las características financieras y demográficas que discriminan las microempresas viables de las que no lo son. Para validar el algoritmo obtenido y ver la precisión predictiva hemos efectuado la clasificación de las 2111 empresas que nos habíamos reservado como muestra de prueba. Asimismo, aplicamos

la técnica de validación cruzada *N-fold*<sup>13</sup>, con el objetivo de reclasificar las microempresas de la muestra de entrenamiento utilizadas para generar las reglas de decisión. Los porcentajes de clasificaciones correctas obtenidas con la aplicación del método *rough set* se presentan en la Tabla 7.

Tabla 7. Indicadores de desempeño de los modelos rough set

Set de variables	Clasificación (10-fold cross-validation)				Pronóstico (muestra de prueba)						
	Eficiencia	Precisión	Especificidad	Sensibilidad	Eficiencia	Precisión	Especificidad	Sensibilidad			
Todos Atributos	68.24	2.61	66.58	63.26	4.18	73.22	4.06	63.05	18.97	63.01	63.31
Reducto 1	66.81	2.57	67.10	67.68	4.67	65.96	4.22	66.16	19.04	67.58	55.76
Reducto 2	61.68	2.57	67.10	59.97	4.49	63.37	7.13	60.23	17.79	59.76	63.77
Reducto 3	66.10	4.94	65.83	65.23	5.64	66.95	7.84	61.19	18.47	61.34	60.07
Reducto 4	65.60	4.31	64.70	62.54	7.20	68.67	4.22	57.33	16.22	56.86	61.15
Reducto 5	63.18	4.87	64.70	61.69	5.74	64.69	7.26	61.17	18.64	60.48	66.19

Fuente: elaboración propia

Se puede observar en la Tabla 7 que el *reducto 1* consiguió el mejor desempeño en el conjunto de los indicadores sobre la muestra de prueba. De este reducto se obtuvieron 380 reglas (193 para la clase 0 y el resto para la clase 1); todas determinísticas ya que la calidad de clasificación es igual a uno. Para interpretar estas reglas, sólo hemos seleccionado aquellas con mayor fuerza (superior a 2%, equivalente a 14 microempresas) para cada clase de decisión. Así en el Anexo 3 Reglas de Decisión obtenidas del Modelo Rough Set, sólo mostramos 26 de las 380 reglas, las que han sido reacomodadas, para fines de presentación, en una matriz de resultados.

## VI.2. Resultados de la aplicación del método de redes neuronales

En el caso de este estudio se utilizó una red neuronal de tres capas, una capa de entrada, una capa oculta, y una de salida, en que las conexiones sólo se dan entre capas sucesivas (*Multi Layer Perceptron -MLP- feedforward*).

Para determinar la estructura óptima de la red se utilizó un *algoritmo genético* que optimizara los parámetros de la red MLP. En la elección de las

13 En este proceso el conjunto total de objetos de la muestra de entrenamiento es dividido en N subconjuntos (*fold*). Las fases de aprendizaje y clasificación son realizadas N veces, con cada uno de los subconjuntos actuando sucesivamente como conjunto de validación y el resto de las subconjuntos como muestra de aprendizaje. La principal idea detrás de esto es que cada objeto es clasificado una vez y sirve para generar las reglas de decisión N-1 veces. El resultado final de *N-fold cross-validation* estará dado por el promedio de los N test individuales.

variables, el algoritmo genético se aplicó a tres conjuntos distintos de datos. El primero consiste en todas las variables disponibles. El segundo grupo se obtuvo de la aplicación del análisis discriminante que contiene sólo aquellas variables más representativas. El tercer grupo consiste en las variables contenidas en el mejor reducto obtenido de la aplicación de la teoría *rough set* caracterizado por contener el menor número de atributos y con variables no correlacionadas.

Podemos observar en la Tabla 8 los resultados de este tipo de modelos y en la Tabla 9 la comprobación de su potencial al ser testeados sobre los datos de validación en una comparativa de medias y desviaciones estándar versus la muestra de entrenamiento. Para controlar el hecho que el error en el proceso de entrenamiento no es un buen estimador del error cometido con los datos de validación, el proceso iterativo de estimación es interrumpido al observar divergencia entre el error en la muestra de entrenamiento y validación.

La mejor red encontrada fue una red MLP con *backpropagation* de tres capas, cada una con una función de activación distinta (modelo D10). Un 71.9% de los casos de la muestra de prueba fueron clasificados correctamente por este modelo. La expresión del modelo neuronal MLP ganador es la siguiente:

$$y = f \left( \theta_j^o + \sum_{j=1}^{10} w_{kj}^o g \left( \sum_{i=1}^{30} w_{ji}^h x_i + \theta_j^h \right) \right)$$

Donde  $y$  es el valor de salida,  $x$  es el vector de entrada definido por las 30 variables explicativas obtenidas del análisis discriminante;  $w_{kj}^o$  es el vector de ponderaciones o parámetros a estimar (une la capa oculta con la salida) y  $w_{ji}^h$  son las ponderaciones que vinculan las entradas con la salida, también a estimar. En el caso de la capa oculta la función de transferencia  $g(\cdot)$ , posee características de *tangente hiperbólica lineal* y, en el caso de la capa de salida, la función de activación  $f(\cdot)$  es una *tangente hiperbólica*. Para esta última fue necesario escalar la salida para estar [0.1] en el intervalo, ya que de lo contrario el proceso de estimación podría quedar alterado.

Tabla 8. Resultados de los modelos de redes neuronales: muestra de prueba

Nombre del modelo	Desempeño de los modelos en muestra de prueba			
	Eficiencia	Precisión	Especificidad	Sensibilidad
T1	63.8	21.3	62.4	74.5
T2	60.6	20.2	58.3	77.3
T3	66.1	22.2	65.2	72.7
T4	56.4	19.2	52.9	81.7
T5	65.7	21.8	65.0	71.2
T6	57.4	19.2	54.4	79.1
T7	60.9	18.8	60.0	67.6
T8	56.8	19.3	53.4	81.7
T9	68.1	22.7	67.9	69.1
T10	55.9	18.6	52.8	79.1

Tabla 8. Continuación

Nombre del modelo	Desempeño de los modelos en muestra de prueba			
	Eficiencia	Precisión	Especificidad	Sensibilidad
D1	65.5	21.9	64.4	73.0
D2	56.5	19.1	53.2	80.6
D3	55.9	19.0	52.3	81.7
D4	61.2	20.8	58.7	79.5
D5	67.0	21.1	67.4	63.7
D6	61.4	20.1	59.7	74.1
D7	60.9	19.6	59.3	72.3
D8	57.8	19.2	55.0	78.1
D9	61.4	19.6	60.1	71.2
D10	71.9	24.1	73.2	62.2
R1	67.1	20.0	68.4	57.9
R2	64.2	20.0	63.9	66.2
R3	67.6	20.3	69.0	57.9
R4	60.7	18.7	59.7	68.0
R5	68.2	20.8	69.5	58.6
R6	53.4	17.6	49.9	78.4
R7	69.6	19.9	72.3	50.4
R8	66.0	19.5	67.1	58.3
R9	69.5	21.0	71.4	55.8
R10	63.2	19.2	63.0	64.4

Fuente: elaboración propia

Tabla 9. Resultados red neuronal por grupo de variables

Grupo de Variables	Media				Desviación Estándar			
	Eficiencia	Precisión	Especificidad	Sensibilidad	Eficiencia	Precisión	Especificidad	Sensibilidad
<b>Clasificación</b>								
Análisis Discr.	74.46	74.19	<b>73.23</b>	75.67	2.03	3.35	5.50	3.98
Rough Set	71.81	70.56	68.07	75.57	<b>1.34</b>	<b>2.43</b>	5.04	3.68
Todas	<b>75.31</b>	<b>74.43</b>	73.19	<b>77.42</b>	1.84	2.77	<b>4.50</b>	<b>3.20</b>
<b>Predicción</b>								
Análisis Discr.	72.79	75.82	78.04	67.53	1.18	3.30	5.60	<b>4.77</b>
Rough Set	72.05	74.11	75.81	<b>68.27</b>	<b>0.69</b>	<b>2.57</b>	<b>5.05</b>	4.92
Todas	71.76	73.92	75.72	67.79	1.14	2.81	5.18	5.46

Fuente: elaboración propia

Aún así, como muestra la Tabla 8, un modelo hecho utilizando las variables resultado del análisis discriminante previo no siempre será mejor. Desde el punto de vista de una aplicación bancaria para ser utilizada en casos reales, el utilizar completa la base de variables tiene, en promedio, mejores resultados que utilizar las variables escogidas por el *rough set*, o por el análisis discriminante. Asimismo, en el caso de la especificidad y de la sensibilidad también es más eficiente en términos de variabilidad de resultados.

El utilizar un modelo como D10, si bien tiene el mejor resultado, también pertenece al tipo de modelos con mayor desviación estándar en los resultados. Ello se traduce en que, si bien se tiene el mejor resultado, hay una variabilidad tan grande, que implica que sea bastante poco probable que efectivamente tengamos un resultado de ese tipo.

En el análisis final, el utilizar cualquier subconjunto de datos es peor, excepto en el caso de la especificidad, en donde la pérdida es negligible. Por el lado de la eficiencia de la estimación también aparece que utilizar todas las variables entrega resultados con mayor estabilidad, sin embargo en el caso de la eficiencia y precisión, no es así. Ahora bien, el aumento en la media en estas dos medidas es sustancial, versus una pérdida en eficiencia no tan importante, especialmente en el caso de la precisión. Todos estos son indicios de que, a la hora de implantar un sistema de estas características, el limitar las variables a utilizar por la red neuronal implicará disminuir los resultados.

### *VI.3. Resultados de la aplicación del método de regresión logística*

Como se señaló anteriormente, la información introducida al método *logit* consistió en las nueve variables cuantitativas obtenidas del análisis factorial a los que agregamos la totalidad de las variables cualitativas. Para seleccionar las variables que forman parte del modelo y que mejor discriminan entre ambos grupos de microempresas se utilizó el criterio *backward stepwise*. Este proceso es un algoritmo de optimización que, partiendo de un modelo que utiliza completamente el conjunto de todas las variables que se consideran candidatos razonables para aumentar la capacidad de predicción del modelo, va eliminándolas y reintegrándolas una a una en cada iteración, de forma tal que obtiene un modelo con un mayor *Log-likelihood*, si es un modelo de elección discreta como un *logit*, o menor Suma de Errores Cuadráticos, en una regresión lineal. En otras palabras, este algoritmo es capaz de diferenciar entre todas las variables disponibles, aquellas que permiten obtener mayor capacidad predictiva<sup>14</sup>.

---

14 Los resultados se muestran en: Anexo 4. Factores Modelo LOGIT y Anexo 5. Resultados Etapas del Modelo LOGIT.

Los resultados que se exponen en la Tabla 10 corresponden a aquel modelo que mostró una mejor clasificación, tomando en cuenta las muestras de entrenamiento y prueba (Step 10).

Tabla 10. Indicadores de desempeño del modelo de regresión logística

		Eficiencia	Precisión	Especificidad	Sensibilidad
Step 1	Clasificación	73.72%	74.45%	75.21%	72.22%
	Predicción	67.40%	22.51%	67.08%	69.78%
Step 2	Clasificación	73.58%	74.37%	75.21%	71.94%
	Predicción	67.49%	22.49%	67.23%	69.42%
Step 3	Clasificación	73.58%	74.37%	75.21%	71.94%
	Predicción	67.40%	22.44%	67.13%	69.42%
Step 4	Clasificación	73.65%	74.41%	75.21%	72.08%
	Predicción	67.40%	22.44%	67.13%	69.42%
Step 5	Clasificación	73.72%	74.45%	75.21%	72.22%
	Predicción	67.40%	22.38%	67.18%	69.06%
Step 6	Clasificación	73.65%	74.41%	75.21%	72.08%
	Predicción	67.36%	22.29%	67.18%	68.71%
Step 7	Clasificación	73.65%	74.34%	75.07%	72.22%
	Predicción	67.49%	22.49%	67.23%	69.42%
Step 8	Clasificación	73.36%	74.19%	75.07%	71.65%
	Predicción	67.49%	22.49%	67.23%	69.42%
Step 9	Clasificación	73.50%	74.26%	75.07%	71.94%
	Predicción	67.58%	22.55%	67.32%	69.42%
Step 10	Clasificación	73.08%	73.82%	74.64%	71.51%
	Predicción	67.71%	22.69%	67.42%	69.78%
Step 11	Clasificación	73.65%	74.20%	74.79%	72.51%
	Predicción	65.97%	21.80%	65.35%	70.50%
Step 12	Clasificación	73.50%	73.91%	74.36%	72.65%
	Predicción	65.50%	21.72%	64.66%	71.58%
Step 13	Clasificación	73.79%	74.34%	74.93%	72.65%
	Predicción	65.37%	21.65%	64.51%	71.58%

Fuente: elaboración propia



Para el caso de la muestra usada para crear el modelo, en promedio, 73.1 por ciento de los casos fueron correctamente clasificados. Finalmente, se aprecia que 67.7 por ciento de los casos de la muestra de prueba fueron correctamente clasificados por el modelo.

## VII. Discusión de resultados

En esta sección se analizan brevemente algunos de los resultados presentados antes. Específicamente, se examinarán las tablas Tabla 7, Tabla 8, Tabla 9 y Tabla 10.

Para comenzar, se observa que los modelos neuronales muestran un desempeño superior en cuanto a la predicción de empresas insolventes que los modelos estimados vía *rough set* y *logit*. Las tasas de eficiencia totales para la muestra de prueba muestran que el modelo neuronal que logró los mejores resultados (modelo D10) alcanzó un porcentaje de eficiencia del orden del 71.9 por ciento. Ninguno de los modelos estimados a través de las otras metodologías logra esta eficiencia.

Si se examinan los grupos solventes e insolventes por separado, se obtienen algunos resultados interesantes. Por ejemplo, si se quiere obtener un modelo que logre una buena predicción con el grupo de *microempresas insolventes* (es decir, una buena sensibilidad), vemos que los modelos neuronales son dominantes. En particular, los modelos T4, T8 y D3 alcanzan un 81.7 por ciento de sensibilidad (Tabla 8). Este resultado es importante, puesto que de nuestra discusión del marco teórico se desprende que *la mayor parte del trabajo de los intermediarios financieros no consiste en seleccionar buenos proyectos sino en descartar malos*. Como vimos, a medida que aumenta la proporción de deudores cuya probabilidad de insolvencia es alta, la tasa de interés que pagan los solventes debe aumentar. Por esta razón, *la selección estricta favorece a los buenos microempresarios porque les disminuye el costo de los fondos*.

Por otra parte, si el interés radica en predecir microempresas del grupo solvente (especificidad), nuevamente los modelos neuronales son superiores con una tasa de eficiencia del orden del 73.2 por ciento. Sin embargo, aquí la diferencia es menor a la lograda con las demás metodologías. Por los resultados de las tablas, en el caso de la especificidad, el modelo *rough set* muestra un desempeño levemente mejor que la regresión logística. La especificidad de los modelos *rough set* en la muestra de prueba es del orden del 67.6 por ciento (Tabla 7) mientras que la regresión logística alcanza el 67.4 por ciento (Tabla 10).

Si se examina la *eficiencia de clasificación* de los modelos estimados según la teoría *rough set*, se ve que la mejor clasificación (proceso *10-fold crossvalidation*) se logra con aquel modelo que usa todas las variables obtenidas del análisis discriminante con una tasa de eficiencia del 68.24 por ciento y un 2.61 por

ciento de desviación estándar<sup>15</sup>. Si se comparan estos resultados con el análisis *logit* y el análisis discriminante existe una superioridad de estos últimos. La tasa de eficiencia lograda por estos modelos es de 73.1 por ciento y 74.6 por ciento (no mostrado), respectivamente.

Al examinar la eficiencia de clasificación de los modelos estimados por las redes neuronales es más notoria la superioridad de los modelos neuronales. Los mejores resultados de clasificación son obtenidos con los modelos T7 y D9 con valores de eficiencia de 78.6 y 78.8 por ciento, respectivamente. Se ve que estos son muy superiores a los obtenidos por los modelos de la teoría *rough set* y *logit*.

Otro aspecto que es interesante destacar es el de la significancia económica de las *variables explicativas más relevantes*. Para ello se considera los modelos *rough set (reducto 1)*, red neuronal D10 y *logit (step 10)*, que resultaron ser los mejores de cada metodología. Se encontró que las variables que aparecen con más frecuencia en todos ellos son: *sexo, región, estado civil, edad, actividades sociales, formalidad de la microempresa, antigüedad en el sector, periodicidad del ahorro, capital original y capacidad de pago*, y la variable macroeconómica *tasa de política monetaria (TPM)*. Esto indica que son variables muy discriminatorias en la muestra entre microempresas solventes e insolventes. En cuanto a los ratios financieros, aparecen con alguna frecuencia los ratios *Ln ventas, KT/AT, ROA, cobertura de intereses, rotación de activos*, y *Ln KT*. En todos los modelos la variable *renegociación* es la más importante. Esta variable está directamente relacionada con características que reflejan un buen o mal desempeño de la microempresa, ya que refleja una situación en que la microempresa no pudo satisfacer una o más condiciones de sus convenios de deuda. De acuerdo a los resultados, las microempresas que renegocian se tratan, en su mayoría, de microempresas que caerán en insolvencia en cualquier caso.

## VIII. Conclusiones

Este trabajo aporta evidencia que facilitará la predicción de insolvencias en microempresas chilenas, ya que la gran mayoría de los estudios realizados, tanto en el ámbito nacional como internacional, modelan sólo la realidad de empresas que cotizan en bolsa y que son de tamaño mediano o grande.

Metodológicamente hemos desarrollado diversos modelos para probar nuestras hipótesis. Por las características de la información disponible para

---

15 En un sentido estricto, la eficiencia de clasificación lograda por la metodología *rough set* de las microempresas usadas para generar el algoritmo, es del 100 por ciento. Sin embargo, dada la especificación del modelo *rough set* éste se ajusta perfectamente a la muestra de aprendizaje. Así, para aproximar su capacidad de clasificación se usó un proceso de reclasificación de las microempresas usadas para generar las reglas de decisión usando el proceso de validación cruzada antes mencionado.

este colectivo de empresas, hemos sugerido el uso de modelos de inteligencia artificial (redes neuronales y *rough set*) y hemos contrastado sus resultados con un modelo estadístico tradicional (regresión logística). En general, encontramos que los modelos de redes neuronales presentan un mejor desempeño clasificatorio y predictivo, con niveles cercanos al 72 por ciento de eficiencia, mientras que los modelos *rough set* y la regresión logística mostraron aciertos de 66.2 y 67.7 por ciento, respectivamente. Indudablemente que los tres tipos de modelos representan un buen avance en la predicción de la insolvencia empresarial de microempresas respecto a la situación actual.

Los resultados nos permiten concluir que la aplicación de modelos flexibles como las redes neuronales y *rough set*, son eficaces en la elección de microempresarios exitosos, sin necesidad de levantar supuestos ni de información preliminar o adicional de ningún tipo sobre los datos, tales como la distribución de probabilidad en los modelos estadísticos. En relación a la teoría *rough set*, aún cuando el desempeño fue, en general, inferior al logrado por las otras metodologías, consideramos que estos resultados son satisfactorios al encontrar un 66.2 por ciento de capacidad predicativa promedio. Al considerar su rapidez para procesar gran cantidad de información, tanto cualitativa como cuantitativa, y su capacidad para generar reglas de decisión fácilmente comprensibles, constituye una alternativa muy competitiva con las técnicas más eficientes.

Los resultados de las metodologías aquí diseñadas establecen la superioridad del modelo de redes neuronales, en cuanto a la clasificación y predicción de la insolvencia de microempresas en Chile, sobre los modelos de regresión logística y *rough set*. Sin embargo, cabe destacar que este último, a diferencia de las redes neuronales permite explicar el fenómeno, lo que se traduce en la identificación de las variables que mejor explican el tema en cuestión. Para salvar esta dificultad de las redes neuronales hemos usado un análisis de sensibilidad que nos permitió interpretar la importancia relativa de cada *input* sobre cada *output* en el método de redes neuronales.

Ha quedado claro que la introducción de variables cualitativas robustece la capacidad de los modelos. Los resultados han demostrado que en el caso de las microempresas, la información contable y activos tangibles del negocio no son suficientes por si solos para predecir el éxito o fracaso de la microempresa. Los mejores predictores de la probabilidad de pago de estas empresas son la historia del crédito personal de su dueño, las posibilidades de refinanciamiento y sus variables sociodemográficas. Así, los principales factores para medir el riesgo de crédito en una microempresa resultaron ser variables asociadas a la historia crediticia del microempresario, tales como si tuvo o tiene otros créditos bancarios; si participó o no en campañas de renegociación; cuál es su experiencia en su negocio; cuánto patrimonio tiene el dueño (por ejemplo, si

posee vehículo); y a variables sociodemográficas como el sexo, edad y región a la cual pertenece el microempresario.

Un aspecto importante a destacar es la existencia de variables disponibles en la base de datos original que no presentaban variación identificatoria entre las categorías de empresas solventes e insolventes. Evidentemente si todos los valores de la variable son iguales entonces no es posible extraer conocimiento alguno de tal variable. Si la variación es muy leve, no seremos capaces de explicar gran parte de la variación en la variable dependiente. A su vez, se debe considerar que se cuenta con poca información respecto a algunas variables de interés y no fue posible utilizarlas como variable de entrada en los modelos, tales como la variable que agrupa información sobre documentos impagos de DICOM, el monto de las moras comerciales, problemas tributarios y otras variables asociadas al grado de compromiso del microempresario con su actividad.

La presente investigación deja abierta la posibilidad de seguir con trabajos futuros, especialmente en el desarrollo de modelos financieros más sofisticados como los elaborados por el Comité de Supervisión Bancaria de Basilea; por ejemplo, el modelo IRD avanzado para la estimación de la probabilidad de incumplimiento. Finalmente, una tarea relevante para cualquier trabajo futuro es determinar los beneficios económicos que conlleva la utilización de las herramientas predictivas, en este caso, los beneficios tangibles y mensurables al poner en uso el modelo que resultó ganador en el análisis realizado anteriormente.

### Referencias Bibliográficas

- Akerlof, G. (1970). The Market for Lemmons: Quality Uncertainly and the Market Mechanism. *Quarterly Journal of Economics*, 89, 488–500.
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 4(XVIII), 589-609.
- Altman, E. I., Haldeman, R. G. y Narayanan, P. (1977). ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. *Journal of Banking and Finance*, 1(1), 29-51.
- Atiya, A. F. (2001). Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results. *IEEE Transactions on Neural Networks*, 4(12), 929-935.
- Bain, J. (1956). *Barriers to New Competition*. Cambridge Massachusetts: Harvard University Press.
- Basch, M. y Montenegro, C. (1989). Aplicación de Modelos Estadísticos Multivariados a la Predicción de Quiebra de Empresas Latinoamericanas. *Paradigmas en administración*, 14.
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research, supplement*, 71 - 127.
- Berger, A. y Udell, G. (1995). Relationships Lending and Lines of Credits in Small Firm Finance. *Journal of Business*, 68, 351–382.
- Beynon, M. J. y Peel, M. J. (2001). Variable Precision Rough Set Theory and Data Discrimination: An Application to Corporate Failure Prediction. *OMEGA: the International Journal of Management Science*, 29(6), 561–576.
- Bravo, D., Crespi, G. y Gutiérrez, I. (2002). *Desarrollo se Escribe con PyME*. Santiago de Chile: FUNDES.
- Bravo, F. (2008). *Análisis de Estados Financieros: Texto y Casos*. Santiago de Chile: Serie Economía - Administración, Editorial Andrés Bello.

- Cabrera, A., De la Cuadra, S., Galetovic, A. y Sanhueza, R. (2002). *Las PyMEs: Quiénes son, Cómo son y Qué hacer con ellas*. Santiago de Chile: Sociedad de Fomento Fabril.
- Chile Emprende, SERCOTEC. (2005). *La Situación de la Micro y Pequeña Empresa en Chile*. Santiago, Chile.
- Crespi, G. (2003). *PyME en Chile: Nace, Crece y...Muere: Análisis de sus Desarrollo en los Últimos Siete Años*. Santiago de Chile: FUNDES.
- Díaz, D. (2005). *Prediciendo el Cambio en el Precio del Oro: Modelos de Redes Neuronales Rolling y Recursivos*. Tesis para optar al título de Magíster en Finanzas, Facultad de Economía y Negocios, Universidad de Chile, Santiago, Chile.
- Dimitras, A. I., Slowinski, R., Susmaga, R. y Zopounidis, C. (1999). Business failure prediction using rough sets. *European Journal of Operational Research*, 114, 263–280.
- Fayyad, U. y Irani, K., (1993). Multi-interval discretization of continuous-valued attributes for classification learning. En *Proceedings of the 13th International Joint Conference on Artificial Intelligence*. Chambéry, France: Kaufmann.
- Fayyad, U. y Irani, K. B. (1992). On the handling of continuous- valued attributes in decision tree generation. *Machine Learning*, 8, 87–102.
- Fayyad, U., Piasteky-Shaphiro, G y Smith, P. (1996). Knowledge Discovery and Data Mining: Towards a Unifying Framework. *Proceeding of the Second International Conference on Knowledge Discovery an Data Mining* (pp. 82- 88).
- FOCUS. (2005). *Desempeño Económico y Viabilidad de la Microempresa en Chile*. Estudio realizado para Fundación Ford. Santiago de Chile.
- Fuentes, R. y Maquieira, C. (2001). Why Borrowers Repay: Understanding High Performance in Chile's Financial Market. En M. Pagano (Ed.). *Defusing Default: Incentives and Instituitions*. Washington: Banco Interamericano de Desarrollo.
- Gujarati, D. (2004). *Basic Econometrics* (4 Ed.). London: McGraw-Hill.
- Greco, S., Matarazzo, B. y Slowinski, R. (1998). A new rough set approach to evaluation of bankruptcy risk. En C. Zopounidis (ed.), *New Operational Tools in the Management of Financial Risks* (pp. 121-136). Dordrecht: Kluwer Academic Publishers.
- Highfield, R. y Smiley, R. (1987). New Business Starts and Economics Activity: An Empirical Investigation. *Internacional Journal of Industrial Organization*, 5.
- Hill, T., Marquez, L., O'Connor, M. y Remus, W. (1994). Artificial neural network models for forecasting and decision making. *International Journal of Forecasting*, 1(10), 5-15.
- Hopenhayn, H. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica*, 5(60), 1127-1150.
- Hornik, K., Stinchcombe M. y White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359-366.
- Jensen, M. y Meckling, W. (1976). Theory of the firm: managerial behavior, agency cost and ownership structure. *Journal of Financial Economic*, 3(4), 305-360.
- Jovanovic, B. (1982). Selection and Evolution of Industry. *Econometrica*, 3(50), 649-670.
- Komorowski, J., Pawlak, Z., Polkowski, L. y Skowron, A. (1999). Rough sets: A tutorial. En S.K. Pal y A. Skowron (Eds.). *Rough Fuzzy Hybridization: A New Trend in Decisionmaking* (pp. 3–98). Singapore: Springer.
- Malhotra, N. K. (2004). *Investigación de Mercados: Un Enfoque Aplicado* (4 Ed.). México: Pearson educación.
- Maltz, A. C., Shenhar, A. J. y Reilly, R. R. (2003). Beyond the balanced scorecard: refining the search for organizational success measures. *Long Range Planning*, (36), 187-204.
- McKee, T. E. (2000). Developing a Bankruptcy Prediction Model via Rough Sets Theory. *International Journal of Intelligent Systems in Accounting, Finance and Management*, 9, 159-173.

- Means, A. (2005). *Modelo de Predicción de Riesgo de las Carteras de Garantía de Crédito del FOGAPE*. Tesis para optar al título de Magíster en Gestión de Operaciones, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.
- Moody, J. E. (1994). Prediction Risk and Architecture Selection for Neural Network. En: V. Cherkassky, J.H. Riedman y H. Wechsler. *Theory and Pattern Recognition Application* (pp.13-26).
- Mosqueda, R. (2005). *Valoración de los Indicadores del Fracaso en las Empresas Mexicanas*. México: Investigación IMEF.
- Myers, S. y Majluf, N. (1984). Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have. *Journal of Financial Economics*, 13, 187-221.
- Nurmi, H., Kacprzyk, J. y Fedrizzi, M. (1996). Probabilistic, fuzzy and rough concepts in social choice. *European journal of Operational Research*, 95, 264-277.
- Odom, M. D. y Sharda, R. (1990). A neural network model for bankruptcy prediction. *Proceedings of the IEEE International Conference on Neural Networks* (pp. 163-168).
- Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 1(18).
- Parisi, A., Parisi, F. y Guerrero, J. L. (2003). Evaluación de Modelos de Redes Neuronales de Predicción del Signo de la Variación del IPSA. Departamento de Administración, Universidad de Chile. pp. 8-9.
- Pawlak, Z. (2002). Rough sets, decision algorithms and Bayes' theorem. *European Journal of Operational Research*, 136(1), 181-189.
- Pawlak, Z. (1997). *Rough Sets and their Applications*. Institute of Theoretical and Applied Informatics Polish Academy of Sciences.
- Pawlak, Z. (1994). Rough set approach to multi-attribute decision analysis. *European Journal of Operational research*, 72, 443-459.
- Pawlak, Z. (1991). *Rough Sets. Theoretical Aspects of Reasoning about Data*. Dordrecht/ Boston/ London: Kluwer Academic Publishers.
- Pawlak, Z. (1984). Rough classification. *International Journal of Man-Machine Studies*, 20, 469-483.
- Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences*, 11(5), 341-356.
- Pawlak, Z., Grzymala-Busse, J., Slowinski, R. y Ziarko, W. (1995). Rough sets. *Communications of the ACM Association for Computing Machinery*, 38(11), 89-97.
- Platt, H. D. y Platt, M. B. (1990). Development of a class of stable predictive variables: the case of bankruptcy prediction. *Journal of Business Finance & Accounting*, 1(17), 31- 51.
- Platt, H. D., Platt, M. B. y Pedersen, J. G. (1994). Bankruptcy Discrimination With Real Variables. *Journal of Business Finance and Accounting*, 21(4).
- Predki, B., Slowinski, R., Stefanowski, J., Susmaga, R. y Wilk, S. (1998). ROSE – Software Implementation of the Rough Set Theory. En L. Polkowski y A. Skowron (Eds.). *Rough Sets and Current Trends in Computing. Lecture Notes in Artificial Intelligence, vol. 1424* (pp. 605-608). Berlin: Springer-Verlag.
- Predki, B. y Wilk, S. (1999). Rough Set Based Data Exploration Using ROSE System. En: Z.W. Ras y A. Skowron (Eds.). *Foundations of Intelligent Systems. Lecture Notes in Artificial Intelligence, vol. 1609*, (pp. 172-180). Berlin: Springer-Verlag.
- Rahimian E., Singh, S., Thammachote, T. y Virmani, R. (1993). Bankruptcy prediction by neural network. En: R. R. Trippi y E. Turban (Eds.). *Neural networks in finance and investing* (pp. 159-171). Probus Publishing.
- Román, E. (2003). Acceso al Crédito Bancario de las Microempresas Chilenas: Lecciones de la Década de los Noventa. *Serie de financiamiento del desarrollo*, 138. Santiago de Chile: CEPAL.
- Romani, G., y otros. (1999). Modelos de clasificación y predicción de quiebra de empresas: una

- aplicación a empresas chilenas. Extraído de <http://upracd.upr.clu.edu:9090/~cicia/Revista/Art2-7-1.pdf>
- Rosembaum, D. y Lamort, F. (1992). Entry barriers, exit and sunk cost: an analysis. *Applied Economics*, 24.
- Sanzana, G. (2006). *Modelo de Predicción de Cierre Voluntario de Cuentas Corrientes en un Institución Bancaria*. Tesis para optar al título de Ingeniería Civil Industrial, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.
- Sanchis, A., Segovia, M. J. y otros (2006). Rough Sets and the role of the monetary policy in financial stability (macroeconomic problem) and the prediction of insolvency in insurance sector (microeconomic problem). *European Journal of Operational Research* (article in press).
- Skowron, A. (1993). Boolean reasoning for decision rules generation. En J. Komorowski, Z. W. Ras. (Eds.). *Methodologies for Intelligent Systems. Lecture Notes in Artificial Intelligence*, vol. 689 (pp. 295–305). Berlin: Springer.
- Skowron, A. (1991). The Implementation of Algorithms based on Discernibility Matrix. Manuscript.
- Smith, R. y Winakor, A. (1935). *Changes in Financial Structures of Unsuccessful Corporations*. Universidad de Illinois.
- Solow, R. (1960). Investment and Technological Progress. En K. Arrow, S Karlin y P. Suppes (Eds.). *Mathematical Methods in the Social Science 1959*. Stanford: Stanford University Press.
- Stiglitz, J. y Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economics Review*, 71, 93-410.
- Wilson, R. L. y Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision Support Systems*, 11(5), 545-557.
- Wroblewski, J. (1998). Genetic algorithm in decomposition and classification problems. En L. Polkowski y A. Skowron (Eds.). *Rough Sets in Knowledge Discovery*, vol. 2 (pp. 471–487). Wurzburg: Physica- Verlag.
- Wroblewski, J. (1995). Finding minimal reducts using genetic algorithm. En Wang, P.P. (Ed.). *Proceedings of the International Workshop on Rough Sets Soft Computing at 2<sup>nd</sup> Annual Joint Conference on Information Sciences (JCIS'95)* (pp. 186–189). Wrightsville Beach, NC.
- Yamawaky, H. (1991). The Effects of Business Condition on Net Entry: Evidence from Japan. En P. Geroski y J. Schwañbach (Ed.). *Entry and Market Contestability: An International Comparison*. Oxford: Basil Blackwell.
- Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: A logistic analysis. *Journal of Business Finance and Accounting*, 1(12), 19-45.
- Zavgren, C. V. (1983). The prediction of corporate failure: the state of the art. *Journal of Accounting Literature*, 1, 1-38.
- Ziarko, W. (1993). Variable precision rough set model. *Journal of Computer and System Sciences*, 46, 39-59.

## Anexos

### *Anexo 1. Medidas de Evaluación para los Modelos.*

Para medir el real desempeño de las metodologías se utilizaron los índices de eficiencia, precisión, especificidad y sensibilidad que se derivan de los cuadrantes de la llamada matriz de confusión. Esta matriz entrega información de los errores y aciertos de la clasificación, y cada cuadrante significa lo siguiente:

		Predicción	
		No Fracaso	Fracaso
Observado	No Fracaso	A	B
	Fracaso	C	D

- A: Predicción acertada, valor predicho es No fracaso y observado es No fracaso.
- B: redicción errónea, valor predicho es fracaso cuando el observado es No fracaso.
- C: Predicción errónea, valor predicho es No fracaso cuando el observado es fracaso.
- D: Predicción acertada, valor predicho es fracaso y observado es fracaso.

El indicador de eficiencia expresa el porcentaje de aciertos que se logra, es el cociente entre los casos que predice acertadamente y el total de casos:

$$= \frac{A+D}{A+B+C+D} \text{ Eficiencia}$$

Por su parte, la precisión indica qué porcentaje de fracasados se consigue reconocer del total de estos. El índice de precisión se define como sigue:

$$= \frac{D}{B+D} \text{ Precisión}$$

Finalmente, los indicadores de especificidad y sensibilidad están definidos por los errores tipo I y tipo II. La especificidad representan la proporción de empresas sanas correctamente clasificadas, mientras que la sensibilidad la proporción de empresas fracasadas correctamente clasificadas.

$$\text{Especificidad} = \frac{A}{A+B}$$

$$\text{Sensibilidad} = \frac{D}{C+D}$$



## Anexo 2. Creación de Ratios Financieros

Con el objetivo de potenciar la predicción de los modelos alternativos, se usaron variables que, de acuerdo a nuestro criterio, entregan mayor información del comportamiento de la microempresa. Estos son ratios financieros clasificados tradicionalmente en la literatura en cinco categorías: liquidez, rentabilidad, endeudamiento o *leverage*, solvencia y eficiencia (ver Beaver, 1967).

Variable creada	Descripción	Función de creación
Tamaño	Es una medida de la escala de la microempresa	$= \ln(\text{ventas anuales})$ donde, $\text{ventas anuales} = \text{ventas meses altos} * \text{número de meses altos} + \text{ventas meses medios} * \text{número de meses medios} + \text{ventas meses bajos} * \text{número de meses bajos}$
ROA	Rentabilidad sobre los activos, variable que mide la eficiencia del dueño para manejar los recursos bajo su control.	$= \text{Resultado operacional} / \text{activo total}$ donde, $\text{resultado operacional} = \text{ventas anuales} - (\text{promedio capacidad de pago mensual} + \text{deuda del negocio} + \text{gastos familiares} - \text{otros ingresos} - \text{ingresos segunda microempresa}) * 12$ $\text{Activo total} = \text{total activo circulante} + \text{total activo fijo productivo} + \text{total activo fijo no productivo} + \text{bien raíz (valor comercial)}$
Rotación de activos	Indica qué tan eficientemente está utilizando el microempresario su inversión en activos totales para generar ventas, o sea, cuántas veces puede colocar entre sus clientes un valor igual a la inversión que se ha realizado en la microempresa.	$= \text{Ventas anuales} / \text{activo total}$
Resultado operacional / ventas	Medida de la capacidad del microempresario para controlar costos	$= \text{Resultado operacional} / \text{ventas anuales}$
Razón Corriente	Medida de la capacidad de pago de deudas de corto plazo	$= \text{Activo circulante} / \text{pasivo circulante}$
Endeudamiento	Es el porcentaje de fondos totales que han sido proporcionado por los acreedores, ya sea en el corto o largo plazo, para invertir en activos, es decir, el grado al cual el microempresario ha financiado sus activos con capital externo.	$= \text{Pasivo total} / \text{activo total}$ Donde, $\text{Pasivo total} = \text{total pasivo circulante} + \text{Total deuda del negocio (saldo)} + \text{total deuda hipotecaria (saldo)} + \text{total deuda largo plazo}$

Bravo - Pinto: Modelos predictivos de la probabilidad...

Variable creada	Descripción	Función de creación
Cobertura de intereses	Indica la capacidad de la microempresa para cumplir con sus obligaciones de pago de intereses	= Resultado operacional / intereses donde, intereses = parte correspondiente a interés de la cuota mensual * 12
Ventas/deuda LP	Medida de eficiencia, indica la capacidad para generar ventas con los recursos financiados por terceros a largo plazo.	= ventas anuales / deuda de largo plazo
Capital de trabajo	Es la diferencia entre el activo circulante y el pasivo circulante, es una medida de liquidez que indica la capacidad de la microempresa para generar fondos suficientes para satisfacer las necesidades corrientes de operación y pagar oportunamente las deudas corrientes.	= Ln (capital de trabajo) donde, capital de trabajo = activo circulante – pasivo circulante
Capital de trabajo / Activos totales	Medida de liquidez.	= Capital de trabajo / Activos totales
Capital de trabajo / Pasivos totales	Medida de liquidez.	= Capital de trabajo / Pasivos totales
Capital de trabajo / ventas	Medida de liquidez.	= Capital de trabajo / ventas
Cuota / Capacidad de pago	Medida de liquidez, indica la capacidad del microempresario para poder cumplir con sus obligaciones mensuales con el banco.	= Cuota mensual / capacidad de pago mensual
IMACEC, desempleo, tasa interés	VARIABLES relacionadas con el estado del ciclo económico que definen el entorno macroeconómico	Estas variables fueron asociadas a cada microempresa con base en la variación experimentada por la variable el mes en que la empresa cae en la categoría fracaso. Para las que no caen en fracaso cuando su morosidad fue máxima. Para las que no son poseen morosidad se considera la fecha de última cuota.

### Anexo 3. Reglas de Decisión obtenidas del Modelo Rough Set

Regla número	Región	Sexo	Edad	Estado civil	Residencia	Renegociación	Capacidad de pago	Ln ventas	Cobertura de i	ROA	Antigüedad en sector	Formalidad	Sercotec	TPM	Clase de decisión	Strength
1		0		2	1				1				0	0	24	(3,42%)
2		1			1	0	1			1	2		0	0	0	18 (2,56%)
3		0		2	1	0		3					0	0	0	32 (4,56%)
4			3				1		1	1					0	18 (2,56%)
5					1	0		2		1	2				0	17 (2,42%)
6					1	0					3		0	0	0	20 (2,85%)
7			1		1	0		3		2				0	0	17 (2,42%)
8				2						1	2	6			0	16 (2,28%)
9						0				1	2	6			0	23 (3,28%)
10	4			2	1										0	17 (2,42%)
11			2			0			3	3			1		0	19 (2,71%)
12	13				1	0	1					3			0	15 (2,14%)
13		0	2	2		0			2						0	19 (2,71%)
14		1		1			1	0	3	3				0	1	33 (4,70%)
15	13									3	0	8	1		1	21 (2,99%)
16	5	1	1		2				3					0	1	24 (3,42%)
17					2				3		0	7		0	1	21 (2,99%)
18	13	1		1	2								1	0	1	29 (4,13%)
19				1			1		3	3	1			0	1	15 (2,14%)
20	13		1		3						0				1	15 (2,14%)
21		0				1	1					1			1	28 (3,99%)
22		1				1					2	8			1	27 (3,85%)
23			1		2	1			2						1	18 (2,56%)
24	13					1							0		1	27 (3,85%)
25		0				1		0					1		1	30 (4,27%)
26		1			1	1									1	21 (2,99%)

Región: 1 (Primera Región), ... , 13 (Región Metropolitana); Sexo: 1 (Masculino) 0 (Femenino); Edad: 0 (-∞,31), 1 [31,53), 2 [53,59), 3 [59,+∞); Estado Civil: 1 (soltero) 2 (casado) 3 (viudo); Residencia 1 (Vivienda Propia) 2 (Arrendatario) 3 (Allegado); Renegociación: 0 (No) 1 (Si); Capacidad de Pago Mensual en UFs: 0 (-∞,3.5) 1 [3.5, 13.1) 2 [13.1, 14.5) 3 [14.5,+∞)[13.1,14.5) 3 [14.5,+∞); LN Ventas en UF: 0 (-∞,6.295) 1 [6.295,6.305) 2 [6.305,6.635) 3 [6.635,+∞); Cob. Intereses en Ufs: 0 (-∞,2.825) 1 [2.825,27.335) 2 [27.335,45.165) 3 [45.165,+∞); ROA en UF: 0 (-∞,0.235) 1 [0.235,0.335) 2 [0.335,0.595) 3 [0.595,+∞); Antigüedad en el Sector en Años: 0 (-∞,4) 1[4,6) 2 [6,22) 3 [22,+∞); Formalidad: 1 (Inicia Actividades,

con Patente Comercial, Cuentas Declaradas registradas en el SII mayores al 50%) 0 (No presenta ninguno de los anteriores); Subsidio Sercotec: 0 (No) 1 (Si); TPM: 0  $(-\infty, 2.125)$  1  $[2.125, 2.375)$  2  $[2.375, 3.125)$  3  $[3.125, +\infty)$  ; Clase de Decisión 0 (Posee Capacidad de Pago) 1 (No posee Capacidad de Pago).

Strenght: Número de objetos que son cubiertos por la regla, y que pertenecen a la clase de decisión sugerida por ella.

#### Anexo 4. Factores Modelo LOGIT seleccionado.

Los resultados que se exponen en la tabla corresponden a aquel modelo que mostró una mejor clasificación (*step 10*), tomando en cuenta las muestras de entrenamiento y prueba.

Variable <sup>[21]</sup>	Beta	Error estándar	Test de Wald	Sig.	Exp (Beta)
<u>Edad</u>	-0.029	0.007	19.575	0.000	0.971
<u>Región</u>			30.416	0.001	
Segunda Región	-1,088	0,610	3,183	0,074	0,337
Tercera Región	-0,852	0,592	2,069	0,150	0,427
Cuarta Región	-1,138	0,479	5,658	0,017	0,320
Quinta Región	-0,441	0,355	1,545	0,214	0,643
Sexta Región	-0,961	0,409	5,505	0,019	0,383
Séptima Región	-0,053	0,504	0,011	0,916	0,948
Octava Región	-0,185	0,414	0,200	0,655	0,831
Novena Región	-0,371	0,510	0,527	0,468	0,690
Décima Región	-0,464	0,403	1,327	0,249	0,629
Duodécima Región	0,519	0,657	0,625	0,429	1,680
Región Metropolitana	0,098	0,334	0,087	0,768	1,103
<u>Sexo (Hombre)</u>	0.620	0.142	19.028	0.000	1.158
<u>Estado civil</u>			6.839	0.033	
Casado	-0,391	0,150	6,749	0,009	0,677
Viudo	-0,133	0,458	0,084	0,772	0,876
<u>Nivel educacional</u>			11.174	0.048	
Básica	-0,619	0,306	4,091	0,043	0,538
Media	0,078	0,184	0,181	0,671	1,081
Técnico-Profesional	0,105	0,409	0,066	0,798	1,110
Superior	-1,418	0,745	3,627	0,057	0,242
Otro	0,202	0,205	0,974	0,324	1,224
<u>Residencia microempresario</u>			54.130	0.000	
Arrendatario	1,154	0,159	52,576	0,000	3,171
Allegado	1,079	0,271	15,817	0,000	2,942
<u>Renegociación</u>	3.823	0.465	67.486	0.000	45.729
<u>Numero de cuotas</u>	0.035	0.025	1.938	0.164	1.035
<u>Asesoría</u>	- 0.538	0.339	2.526	0.112	0.584
<u>Tipo Garantía</u>	- 0.438	0.271	2.622	0.105	0.645
<u>Actividades sociales</u>	-0.727	0.157	21.494	0.000	0.483
<u>Estudio de Hijos</u>			7.244	0.027	
Estudio de Hijos(Si)	-0,148	0,309	0,230	0,631	0,862
Estudio de Hijos(No)	0,359	0,152	5,594	0,018	1,432

### Anexo 4. Continuación

<b>Vehículo</b>	- 0.654	0.179	13.373	0.000	0.520
<b>Tuvo o Tiene créditos</b>	- 1.255	0.174	51.788	0.000	0.285
<b>Formalidad microempresa</b>			19.925	0.006	
<b>Perfil 2</b>	-0,017	0,413	0,002	0,967	0,983
<b>Perfil 3</b>	0,444	0,422	1,108	0,292	1,559
<b>Perfil 4</b>	-0,396	0,591	0,449	0,503	0,673
<b>Perfil 5</b>	0,181	0,430	0,177	0,674	1,198
<b>Perfil 6</b>	-0,384	0,414	0,858	0,354	0,681
<b>Perfil 7</b>	0,358	0,426	0,708	0,400	1,431
<b>Perfil 8</b>	0,471	0,381	1,524	0,217	1,601
<b>Cobertura intereses</b>	0.005	0.002	6.363	0.012	1.005
<b>Ln (KT)</b>	-0.100	0.030	10.792	0.001	0.905
<b>Tasa interés (TPM)</b>	-0.394	0.221	3.166	0.075	0.675
<b>Constante</b>	1.942	0.838	5.371	0.020	6.975
<b>2 Log Likelihood: 1400,613</b>	Hosmer-Lemeshow (sig.) :0.251				
<b>Cox &amp; Snell R Square: 0.322</b>	Nagelkerke R Square:0.429				

### Anexo 5. Resultados Etapas del Modelo LOGIT

Observado			Predicción			
			Muestra de entrenamiento (a)		Muestra de prueba (b)	
			Fracaso		Fracaso	
		Microempresa no fracasada	Microempresa fracasada	Microempresa no fracasada	Microempresa fracasada	
Step 1	Fracaso	Microempresa no fracasada	528	174	1.361	668
		Microempresa fracasada	195	507	84	194
Step 2	Fracaso	Microempresa no fracasada	528	174	1.364	665
		Microempresa fracasada	197	505	85	193
Step 3	Fracaso	Microempresa no fracasada	528	174	1.362	667
		Microempresa fracasada	197	505	85	193
Step 4	Fracaso	Microempresa no fracasada	528	174	1.362	667
		Microempresa fracasada	196	506	85	193

Observado		Predicción				
		Muestra de entrenamiento (a)		Muestra de prueba (b)		
		Fracaso		Fracaso		
		Microempresa no fracasada	Microempresa fracasada	Microempresa no fracasada	Microempresa fracasada	
Step 5	Fracaso	Microempresa no fracasada	528	174	1.363	666
		Microempresa fracasada	195	507	86	192
Step 6	Fracaso	Microempresa no fracasada	528	174	1.363	666
		Microempresa fracasada	196	506	87	191
Step 7	Fracaso	Microempresa no fracasada	527	175	1.364	665
		Microempresa fracasada	195	507	85	193
Step 8	Fracaso	Microempresa no fracasada	527	175	1.364	665
		Microempresa fracasada	199	503	85	193
Step 9	Fracaso	Microempresa no fracasada	527	175	1.366	663
		Microempresa fracasada	197	505	85	193
Step 10	Fracaso	Microempresa no fracasada	524	178	1.368	661
		Microempresa fracasada	200	502	84	194
Step 11	Fracaso	Microempresa no fracasada	525	177	1.326	703
		Microempresa fracasada	193	509	82	196
Step 12	Fracaso	Microempresa no fracasada	522	180	1.312	717
		Microempresa fracasada	192	510	79	199
Step 13	Fracaso	Microempresa no fracasada	526	176	1.309	720
		Microempresa fracasada	192	510	79	199

a. Selected cases validation EQ 0

b. Unselected cases validation NE 0

d. The cut value is .500

## Anexo 6. Resultado Muestra de Entrenamiento Redes Neuronales

Set de Variables	Nombre modelo	Función de activación		Arquitectura del modelo		Desempeño de los modelos en muestra de entrenamiento		Desempeño de los modelos en muestra de validación										
		Oculto	Salida	Neuronas oculta	Neuronas oculta	Tasa de aprendizaje	Tasa de momentum	Épocas	Clasificación (muestra de entrenamiento)	Predicción (muestra de validación)								
				Oculto	Salida	Oculto	Salida	Efic.	Espec.	Sens.	Efic.	Espec.	Sens.					
Todas	T1	Sigmoideal	Sigmoideal	46	46	0.12	0.7	0.9	1000	73.9	74.5	75.2	72.6	72.4	77.3	81.4	63.3	
Todas	T2	Sigmoideal	Tangente	34	34	0.3	0.89	0.82	0.74	1000	73.4	70.9	67.5	79.3	74.0	74.2	74.3	73.8
Todas	T3	Sigmoideal	Lineal	12	12	0.39	0.81	0.51	0.53	1000	74.9	75.7	76.4	73.4	71.9	77.4	81.9	61.9
Todas	T4	Sigmoideal	Lineal Tan. H.	14	14	0.86	0.16	0.98	0.57	1000	75.6	75.7	75.8	75.4	73.1	70.8	67.7	78.5
Todas	T5	Tangente Hip.	Sigmoideal	15	15	0.23	0.41	0.2	0.35	1000	77.3	77.8	78.3	76.4	71.4	76.5	81.0	61.9
Todas	T6	Tangente Hip.	Tangente	19	19	0.44	0.78	0.85	0.31	1000	74.3	71.2	66.9	81.7	70.0	69.6	69.0	71.0
Todas	T7	Tangente Hip.	Hip.	13	13	0.46	0.68	0.9	0.55	1000	78.6	77.8	77.2	79.9	72.1	74.6	77.1	67.1
Todas	T8	Tangente Hip.	Lineal	18	18	0.19	0.26	0.68	0.99	1000	72.4	70.8	68.5	76.2	71.0	71.4	71.9	70.0
Todas	T9	Lineal Tan. H	H.	15	15	0.11	0.17	0.43	0.95	1000	77.2	77.6	77.8	76.6	71.2	76.3	81.0	61.4
Todas	T10	Lineal Tan. H.	Sigmoideal	88	14	0.85	0.26	0.29	0.05	1000	75.5	72.3	68.3	82.7	70.5	71.1	71.9	69.0
AD	D1	Sigmoideal	Hip.	19	19	0.25	0.59	0.77	0.11	1000	74.3	75.5	76.6	72.0	73.8	79.1	82.9	64.8
AD	D2	Sigmoideal	Tangente	55	32	0.27	0.54	0.25	0.82	1000	72.4	70.4	67.2	77.7	73.8	73.6	73.3	74.3
AD	D3	Sigmoideal	Hip.	55	12	0.86	0.37	0.84	0.55	1000	70.9	70.3	69.0	72.8	73.3	75.8	78.1	68.6
AD	D4	Sigmoideal	Lineal Tan.	55	16	0.25	0.72	0.23	0.37	1000	73.3	73.5	73.5	73.0	72.9	76.7	80.0	65.7
AD	D5	Tangente Hip.	H.	55	11	0.21	0.01	0.91	0.14	1000	75.3	76.2	76.8	73.8	73.3	77.5	81.0	65.7
AD	D6	Tangente Hip.	Sigmoideal	55	11	0.18	0.08	0.27	0.29	1000	75.0	72.9	70.3	79.7	71.0	72.0	73.3	68.6
AD	D7	Tangente Hip.	Hip.	55	14	0.63	0.49	0.92	0.08	1000	74.6	74.6	74.5	74.6	74.8	79.9	83.3	66.2
AD	D8	Tangente Hip.	Lineal Tan.	55	15	0.36	0.81	0.12	0.6	1000	73.9	69.7	63.1	84.6	71.2	68.9	65.2	77.1
AD	D9	Lineal Tan. H.	H.	55	29	0.97	0.47	0.36	0.49	1000	78.8	79.6	80.0	77.5	71.9	75.6	79.0	64.8
AD	D10	Lineal Tan. H	Tangente	55	10	0.77	0.39	0.9	0.34	1000	76.1	79.2	81.3	71.0	71.9	79.1	84.3	59.5
RST	R1	Sigmoideal	Hip.	36	38	0.12	0.7	0.73	0.4	1000	72.7	72.2	71.5	73.8	71.7	74.9	78.1	65.2
RST	R2	Sigmoideal	Tangente	36	18	0.72	0.31	0.6	0.38	1000	71.0	70.6	69.7	72.4	72.6	75.4	78.1	67.1
RST	R3	Sigmoideal	Hip.	36	39	0.64	0.15	0.76	0.4	1000	71.6	70.6	68.8	74.4	72.4	74.7	77.1	67.6
RST	R4	Sigmoideal	Lineal Tan.	36	22	0.56	0.44	0.15	0.87	1000	70.4	68.4	64.6	76.3	73.1	73.9	74.8	71.4
RST	R5	Tangente Hip.	H.	36	35	0.36	0.82	0.06	0.11	1000	73.8	74.0	74.1	73.4	72.1	76.3	80.0	64.3
RST	R6	Tangente Hip.	Tangente	36	20	0.86	0.3	0.8	0.84	1000	71.4	67.3	59.3	83.6	71.4	69.2	65.7	77.1
RST	R7	Tangente Hip.	Hip.	36	14	0.93	0.97	0.81	0.28	1000	71.3	69.4	66.0	76.7	71.7	72.4	73.3	70.0
RST	R8	Tangente Hip.	Lineal Tan.	36	16	0.62	0.55	0.35	0.82	1000	71.1	71.6	72.1	70.2	73.1	76.5	79.5	66.7
RST	R9	Lineal Tan. H.	H.	36	23	0.5	0.71	0.73	0.11	1000	74.5	74.4	74.1	74.8	71.0	77.5	82.9	59.0
RST	R10	Lineal Tan. H.	Tangente	36	19	0.81	0.2	0.36	0.8	1000	70.3	67.1	60.5	80.1	71.4	70.3	68.6	74.3
			Hip.															