



Campanaro, Rosa
Díaz, Daniel
Marchese, Alicia
Sepliarsky, Patricia
Viola, Mónica

Riveros, J.
Carrizo, F.
Ferraro, A.
Díaz Toledo, S.
Gardenal, L.

Instituto de Investigaciones y Asistencia Tecnológica en Administración, Escuela de Administración.

MODELOS DE PREDICTIBILIDAD DE QUIEBRAS E INSOLVENCIA BASADOS EN ANÁLISIS DE ESTADOS FINANCIEROS. EVALUACIÓN CRÍTICA Y ASPECTOS METODOLÓGICOS ENFOCADOS EN EL USO DE HERRAMIENTAS DE B.I.

Resumen

El presente trabajo está basado en el Proyecto de Investigación APLICACIÓN DE HERRAMIENTAS INTELIGENTES AL ANÁLISIS DE INFORMES FINANCIEROS". La SEC (Securities and Exchange Commission) de USAEI viene recibiendo desde el 2005, reportes financieros de sus empresas reguladas, codificados con el lenguaje XBRL (eXtensible Business Reporting Language) es decir en formato apto para su explotación por sistemas computacionales. A diciembre 2015 se podía acceder a más de 163.000 presentaciones. Este repositorio de información económico-financiera representa una apreciable fuente de datos para el estudio de modelos analíticos y predictivos de comportamiento empresarial. En el presente trabajo se analiza el modelo de predictibilidad de quiebras de Altman, seleccionado por el proyecto de investigación por su relevancia, a fin de evaluar su eficiencia en base a datos actuales. Se analiza el enfoque tecnológico necesario para la gestión de cantidades masivas de datos, la factibilidad de aplicación de técnicas de inteligencia artificial para el descubrimiento de conocimiento, y por último se contemplan nuevas líneas de investigación fundadas en técnicas de análisis de ratios (Key Performance Indicators).

Palabras claves: Predictibilidad - Informes Financieros - XBRL - Ratios - Data Mining - Big Data

Abstract

The present work is based on Research Project: "USING INTELLIGENT TOOLS FOR ANALYZING FINANCIAL REPORTS". The SEC (Securities and Exchange Commission) from USA is gathering financial reports of public companies since 2005. Those reports are codified with XBRL (eXtensible Business Reporting Language). By this way, they're in a format enable for directly be used in computational system. At december 2015, the system counted with 163.000 filings. This repository of financial and economic information means a value source of data for the study of analytic and predictive models of business. This work analyze the bankruptcy's predictable model of Altman, chosen by the research project, for its relevance, and with the goal of evaluating it efficiency under the scope of actual data. Is being analyzed the technologic approach necessary for managing big volume of Data (Big Data), the feasibility of applying artificial intelligence tools for knowledge discovery, and finally new lines



of research are proposed, based on ratios analyzed techniques (Key Performance Indicators).

Keywords: Predictability - Financial Reports - XBRL - Key Performance Indicators - Data Mining - Big Data

Antecedentes

Objetivos propuestos

Aplicar técnicas de Inteligencia de Negocios - Minería de Datos, a fin de detectar correlaciones entre variables relevantes, inferir patrones de comportamiento y finalmente desarrollar postulaciones de nuevos modelos de análisis a ser evaluados tanto en ámbitos académicos como profesionales.

A partir de los mismos, se espera: Corroborar y evaluar críticamente los modelos que hoy son tomados en la disciplina contable como parámetros de análisis. Proponer un nuevo modelo de análisis de Estados Financieros, consecuencia del estudio de patrones y correlación de datos. Desarrollar una metodología específica para sustentar el grado de certeza asignado al modelo obtenido.

Metodología aplicada

Una vez actualizado el fichado de la bibliografía existente sobre el tema, incluyendo a los autores más reconocidos, sus fundamentos y el marco histórico-tecnológico en que se realizó la producción bajo análisis, se procedió a la apropiación de los Informes Financieros disponibles por la SEC - USA y a la construcción de un almacenamiento en una base de datos "no-SQL". Éste constituye el insumo necesario para la evaluación primitiva, y su posterior "clusterización". Actualmente nos encontramos en la tarea de analizar los informes almacenados y clasificados.

Se trata de una investigación de tipo descriptiva, de características cuali - cuantitativas, cuyo principal enfoque se orienta al análisis de datos. Se jerarquizará la generación de una robusta metodología específica del modelo a desarrollar, con el objeto de sustentar grados de certeza razonablemente válidos de los postulados del mismo. Se espera realizar iteraciones sucesivas, a la luz de los resultados obtenidos con el análisis inteligente y los caminos a tomar ante cada respuesta del sistema.

Introducción

En el mundo de los negocios siempre ha sido sustancial contar con herramientas que permitan medir el riesgo financiero de las empresas altamente competitivas dentro de cada mercado y bajo un contexto como ya sabemos globalizado. La información que se pueda suministrar sobre los riesgos financieros de empresas, será altamente valorada tanto para los dueños de las acciones como para aquellos potenciales inversores, instituciones financieras, proveedores y clientes, e interesados en general, en evaluar la gestión empresarial.

En la primer etapa de este proyecto (2015) nos propusimos profundizar el análisis con una elección de los ratios más adecuados a la tarea y su aplicación al conjunto de datos



obtenidos de la SEC (Securities Exchange Commission), basado en informes presentados a ese organismo a partir de junio 2005, con el objeto de concluir acerca de los más adecuados para la predicción de quiebras o insolvencias.

Una herramienta de evaluación que se va a analizar en el presente trabajo es el modelo "Z-score" desarrollado por Edward Altman, utilizado para medir situaciones de insolvencia financiera, el cual será objeto de estudio en esta parte del proyecto SECBI.

Este modelo "Z" es una técnica de análisis discriminante múltiple. A continuación se desarrollan los fundamentos basales de esta técnica.

Análisis discriminante

El análisis multivariante es un conjunto de métodos estadísticos y matemáticos destinados a describir e interpretar los datos que provienen de la observación de varias variables estadísticas, estudiadas conjuntamente.

Asumiendo que un conjunto de casos de estudio está ya clasificado en una serie de grupos; es decir, se sabe previamente a qué grupos, pertenecen, el análisis discriminante es una técnica del análisis multivariante cuyo objetivo es encontrar la combinación lineal de las variables independientes que mejor permite diferenciar (discriminar) a los grupos.

Una vez encontrada esa combinación (la función discriminante) podrá ser usada para clasificar nuevos casos de estudio. Aprovecha las relaciones existentes entre una gran cantidad de variables independientes para maximizar la capacidad de discriminación.

Beaver, en 1966, utilizando el análisis discriminante simple planteó un modelo de predicción de quiebra. Altman, en 1968, utilizando el análisis discriminante múltiple planteó también un nuevo modelo de predicción de quiebra, el cual ha sido ampliamente difundido.¹

El Z-score de Altman es una versión personalizada de la técnica de análisis discriminante de RA Fisher (1936).

El trabajo de William Beaver, publicado en 1966 y 1968, fue el primero en aplicar un método estadístico, t-test para predecir la quiebra de una muestra de empresas. Beaver aplicó este método para evaluar la importancia de cada uno de varios coeficientes de contabilidad basado en el análisis univariable, utilizando cada relación de la contabilidad de una en una. La principal mejora de Altman fue aplicar un método estadístico, análisis discriminante, lo que podría tener en cuenta múltiples variables simultáneamente.

Debido a la gran cifra de variables que se encuentran y que son importantes indicadores de problemas corporativos en estudios anteriores, se compila una lista de veinte variables potencialmente útiles para la evaluación.

En su estudio Altman como ya se mencionó, seleccionó 22 ratios para ser evaluados, de los cuales se escogieron aquellos que mostraron mayor poder predictivo y que, al mismo tiempo, minimizan la correlación entre sí, es decir, se intentó seleccionar aquellos ratios que concentran la mayor cantidad de información.

¹ Janet Cecibel Aldazábal Contreras -Alberto Fernando Napán Vera Revista de la Facultad de Ciencias Contables Vol. 22 N.º 42 pp. 53-59 (2014) UNMSM, Lima - Perú ISSN: 1560-9103 (versión impresa) / ISSN: 1609-8196 (versión electrónica)



La lista de ratios fue reducida a cinco. El procedimiento que se realizó para discriminar entre el grupo de variables fue el siguiente:

- Observación de la significancia estadística de varias funciones alternativas, analizando también la contribución relativa de cada variable independiente.
- Evaluación de las inter correlaciones entre las variables relevantes.
- Observación de la efectividad predictiva de varios grupos de ratios.
- Evaluación según el criterio del investigador.

Uso del análisis discriminante múltiple (ADM)

El uso del ADM fue propuesto por Altman como modelo predictivo de quiebra de una empresa. El modelo de ADM combina la información de diversas variables financieras en un análisis interdependiente.

El primer paso es determinar los grupos de clasificación, en nuestro caso:

a) Insolvencia - b) Solvencia

Esta predicción de pertenencia o no a un grupo se constituye en la variable dependiente. Establecidos los grupos, se recolectan los datos de las empresas que los componen.

El modelo de ADM permite determinar un conjunto con los coeficientes que discriminan mejor a los grupos de estudio. Estos coeficientes son ponderados y se construye una ecuación lineal con el cual clasificar a la empresa en estudio.

La función discriminante tiene la forma: $Z = V_1 X_1 + V_2 X_2 + \dots + V_n X_n$.

Altman calculó 22 ratios financieros clasificados en 5 categorías estándar:

$$Z = 1.23 X_1 + 1.4 X_2 + 3.3 X_3 + 0.6 X_4 + 0.99 X_5$$

Donde:

X1 = Capital de trabajo/Activo total.

X2 = Utilidades retenidas/Activo total.

X3 = Utilidades antes de intereses e impuestos/Activo total.

X4 = Valor de mercado de las acciones/Pasivo total.

X5 = Ventas/Activo total.²

Según los valores que obtuviera Z al ser evaluado con las cifras de la empresa estudiada, el resultado puede ser:

- a) Saludable: la empresa no tendrá problemas de insolvencia en el futuro;
- b) Zona gris: la empresa tiene altas posibilidades de caer en insolvencia dentro de los próximos 2 años, sugiere reconsiderar estrategias e implementar cambios.
- c) Enferma: la probabilidad de insolvencia en el futuro es muy alta, podría presentar quiebra en 2 años.

² Altman decidió excluirlo del modelo Z2



Representatividad de los coeficientes de la fórmula score “Z”

A continuación se realiza un análisis de los coeficientes seleccionados para la determinación de la fórmula propuesta por el modelo.

Cociente X1 = Capital de trabajo/total de activos

En el análisis financiero, el uso de los activos totales como denominador es habitual, por lo que es lógico que haya ese denominador en cuatro de los cinco ratios.

El activo total en el denominador permite comparar los otros elementos del balance o del estado de resultados.

El numerador es la diferencia entre los activos a corto plazo y pasivos a corto plazo.

Puede ser negativo si los pasivos a corto plazo son mayores que los activos a corto plazo.

Esta relación proporciona información acerca de la situación financiera a corto plazo de la empresa.

Mayor nivel de capital de trabajo en comparación con el total de activos significa mayor nivel de liquidez. Altman distinguió a X1 como una de los ratios más importantes.

El capital de trabajo se define como la diferencia entre activos circulantes y pasivos circulantes. Las características de liquidez y tamaño son específicamente consideradas.

Normalmente, cuando una empresa experimenta pérdidas operativas, tendrá una reducción del activo corriente en relación con el activo total. De los tres ratios de liquidez evaluados, este demostró ser el más valioso.

La inclusión de esta variable es consistente con el estudio de Merwin que califica el capital de trabajo en relación a los activos totales como el mejor indicador de quiebra definitiva.^{3,4}

Cociente X2 = Utilidades retenidas acumuladas / Activos totales (en %)

El cociente X2 = (UR / AT): es uno de los factores más representativos, porque es una medida de largo plazo frente a las políticas de retención de utilidades. A pesar de esto, las mismas pueden ser manipuladas, de alguna manera, y podría crearse algún sesgo en sus resultados.

Esta manipulación podría darse a través de quasi-reorganizaciones empresarias y de declaraciones de dividendos en acciones.

Los beneficios no distribuidos son aquellos que la empresa no reparte a sus accionistas en forma de dividendos, sino que los conserva para ser utilizados dentro de empresa, por ejemplo, para invertirlos en aumentar sus ventas o su producción y hacer crecer sus beneficios futuros.

Las utilidades retenidas o beneficios no distribuidos se calculan restando los dividendos al beneficio neto.

³ Ringelin (2004) citando el estudio de Altman “ Financial ratios, discriminant analysis and the prediction of corporate bankruptcy”

⁴ Altman “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy” (1968)



Su cálculo sería el siguiente:

Beginning retained earnings (Beneficios no distribuidos al inicio) + Net income during the period (Ingresos netos del periodo) - Dividends paid (Dividendos pagados) = Ending retained earnings (Beneficios no distribuidos finales)

Este ratio lleva implícito un doble análisis:

Existe una relación entre la edad de la empresa y este cociente. Por ejemplo, una compañía relativamente joven, probablemente exprese un ratio X2 relativamente bajo porque quizá no ha tenido el tiempo para construir sus beneficios acumulados. Más aún, por este motivo, el ratio antes mencionado puede resultar discriminatorio y jugarle en su contra a este tipo de empresas "jóvenes", llevándolas a ser clasificadas como posibles sujetos de quiebra. Esto, en contraste con empresas de mayor antigüedad, ceteris paribus. Sin embargo, en la práctica esto tiene sentido, puesto que el riesgo de estas empresas iniciadas con un menor plazo de vida suele ser mucho mayor y estar mucho más expuestas a la situación de bancarrota.

Cociente X3 = Utilidades antes de intereses e impuestos/Activo total

Es un ratio cuyo objetivo es medir la verdadera productividad de los activos de una empresa. Toma las ganancias antes de impuestos e intereses y el total de los activos.

Una empresa basa su existencia en la capacidad de obtener ingresos de sus activos.

Altman consideró que esta relación era "particularmente apropiada para los estudios que se ocupan del fracaso empresarial"⁵.

La insolvencia, en su sentido de quiebra, se produce cuando los pasivos excede el valor de los activos medidos, razonablemente, por el valor adquisitivo de los mismos.

Consideró que este ratio supera otras formas de medir la rentabilidad como el flujo de caja.

Tal como define Van Horne el cociente de Utilidad Neta sobre Activo total permite "determinar la eficiencia global en cuanto a la generación de utilidades con Activos disponibles"⁶. Esto nos permite analizar el perfil del poder productivo que tiene el capital invertido.

Cociente X4 = Valor de mercado de las acciones / Pasivo total

Es un ratio muy utilizado al momento de determinar el riesgo crediticio de una empresa, ya que nos muestra las veces en que el valor de liquidación contable de las empresas es capaz de cubrir el total de pasivos. Por lo general, mientras tanto menor que 1 sea este resultado, mejor será la capacidad de endeudamiento.⁷

El X4 muestra hasta dónde puede disminuir el valor del patrimonio, antes de que sea superado por el total de deudas y que, de esta manera, la firma se torne insolvente. Con este

⁵ Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. Edward I. Altman. The Journal of Finance, Vol. 23, No. 4. (Sep., 1968), pp. 589-609.

⁶ Van Horen, J. C. W., & John, M. Fundamentos de administración financiera/por James C. Van Horne y John M. Wachowicz (No. 658.15 V3y 1992).

⁷ Alberto Ibarra Mares "Una perspectiva sobre la evolución en la utilización de las razones financieras o ratios "pensamiento y gestión, n° 21 universidad del norte issn 1657-6276 (Pág.38, 40)



ratio, se incorpora una parte de la información de mercado (el precio de las acciones), pero claramente no queda reflejada la parte más relevante de la información de mercado, a saber, la varianza de los precios de las acciones.

Es la variable con el peso más bajo y el motivo es su numerador: el valor de mercado de las acciones.

Estos datos se obtienen de los mercados de capitales y, eventualmente, puede ser fácil de manipular. El denominador es el pasivo total por lo que la relación muestra si el valor de mercado del patrimonio excede al valor en libros de los pasivos. Existe un riesgo potencial de una empresa si este no es el caso. Entonces la empresa debería encontrar la forma de disminuir sus pasivos.⁸

Este factor (variable independiente X 4) era en realidad el apalancamiento financiero, pues se refería a aquellas operaciones financieras rentables efectuadas con préstamos (relación deuda / capital propio; o la relación endeudamiento / medios propios). Es decir, se trataba de la compra de activos a cambio de emisión de obligaciones.⁹

Para Altman los cinco factores mostraron ser las mejores combinaciones para el discriminante entre empresas en quiebra y empresas sin quiebra.

Sin embargo, este autor no indicó con base a que dividió en cinco categorías su modelo, y si en realidad dichas categorías eran las más representativas en su conjunto e independientes entre sí para predecir una quiebra.

También en su primer trabajo llama la atención que de los veintidós ratios seleccionados no se consideró al ratio de cash flow / deuda total, el cual había proporcionado buenos resultados de clasificación en los estudios de Beaver^{10, 11} y en otros trabajos posteriores.

Según Altman, esto se debió a la dificultad que existió para obtener bases de datos que incluyeran la amortización.

Cociente X5 = Ventas / Activos totales

X5= VN/ AT: este es uno de los indicadores más significativos, porque ilustra las ventas que son generadas por la inversión total de la compañía representada por sus activos.

⁸ Janet Cecibel Aldazábal Contreras -Alberto Fernando Napán Vera Revista de la Facultad de Ciencias Contables Vol. 22 N. ° 42 pp. 53-59 (2014) UNMSM, Lima - Perú ISSN: 1560-9103 (versión impresa) / ISSN: 1609-8196 (versión electrónica)

⁹ Alberto Ibarra Mares- ANÁLISIS DE LAS DIFICULTADES FINANCIERAS DE LAS EMPRESAS EN UNA ECONOMÍA EMERGENTE: LAS BASES DE DATOS Y LAS VARIABLES INDEPENDIENTES EN EL SECTOR HOTELERO DE LA BOLSA MEXICANA DE VALORES -<http://www.eumed.net/tesis-doctorales/2010/aim/MODELO%20ALTMAN.htm>

¹⁰ Beaver (1966) desarrolló su estudio desde un punto de vista univariante lo que, a pesar de su importancia como punto de partida de esta línea de investigación, presenta el inconveniente de no tener en cuenta las posibles relaciones entre variables.

García Gallego Ana, Mures Quintana María Jesús "La muestra de empresas en los modelos de predicción del fracaso: influencia en los resultados de clasificación" Revista de métodos cuantitativos para la economía y la empresa Pág. 135 Junio 2013 ISSN: 18869-516X-<http://www.upo.es/RevMetCuant/pdf/vol15/art73.pdf>



Frente a las críticas del modelo Z original, en las cuales se describía que los modelos anteriores no eran aplicables a las empresas comerciales y de servicio, Altman realizó modificaciones en su estudio y llegó a obtener el valor Z2, para ser aplicada por cualquier tipo de empresa diferente a las contempladas en Z y Z1. Entonces Z2 debe aplicarse cuando la empresa es comercial o de servicios. Esta versión del modelo, fue realizada para corregir la distorsión que podría presentarse, debido a que elimina la razón de rotación de activos (X5) y otorga un peso relativo aún más importante al factor de generación de utilidades en relación al activo (X3). De esta manera en Z2 se eliminó X5 y se mantuvieron los demás índices utilizados en la versión Z1. El argumento para la eliminación del X5 (ventas/activos totales), es que este índice es problemático en el sentido de que varía significativamente de una a otra empresa de distinto sector industrial, lo que trae distorsión y puede dar lugar a interpretaciones erróneas que distorsionan el modelo.

Este ratio permite medir de cierta forma que tan capaces son los administradores de la firma para desenvolverse en situaciones de competencia, ilustrando su habilidad para generar ventas.

Tomando el cociente de forma aislada, este resulta ser el menos significativo. Su implicancia en el modelo surge de su relación con los componentes de los otros ratios.

En un contexto multivariable, toma fuerza por su correlación negativa con el ratio X3. La lógica de esta correlación negativa surge de que, a medida que las firmas sufren pérdidas y se van deteriorando con camino al fracaso, sus activos no son reemplazados o recuperados como en los periodos de "bienestar", y, además, las pérdidas acumuladas van reduciendo aún más el tamaño de los activos a través de débitos a las ganancias retenidas. La pérdida de activos, aparentemente dominan los movimientos de las ventas.^{12,13,14, 15}

Ventajas de utilizar ADM (Análisis Discriminante)

El análisis discriminante múltiple fue seleccionado como el método de análisis estadístico. La primera ventaja de utilizar el ADM sobre el análisis univariable es que el primero analizaba el perfil completo de las características simultáneamente y no sólo individualmente.

La experiencia con este modelo llevó a su autor a la conclusión de que puntuaciones Z-Scores inferiores a 1.81 indicaban una probabilidad elevada de quiebra. En cambio, puntuaciones superiores a 3.00 indicaban una escasa probabilidad de quiebra. Altman denominó "la zona de ignorancia" al rango comprendido entre 1.81 y 2.99.

¹² <http://www.accountingtools.com/>

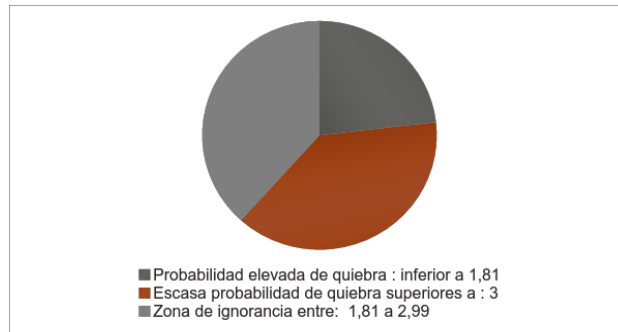
¹³ César A. León Valdés, "El análisis financiero como herramienta en la predicción de quiebra e insolvencia financiera. Universidad Externado de Colombia
<http://revistas.uexternado.edu.co/index.php/contad/article/view/1295/1232>

¹⁴ Ringelin (2004) citando el estudio de Altman "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy".

¹⁵ Altman "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy". (1968)



Figura 1: Puntuación Z-Scores Altman Fuente: a partir de la Investigación 1ECO172



Al analizar a las empresas que cayeron dentro de la zona de ignorancia, Altman razonó que utilizando una Z-Score de 2.675 daba como resultado el número más bajo de errores para clasificar a una empresa entre quiebra y no quiebra. Según él, este punto medio de la Z-Score podía tener aplicaciones prácticas para la evaluación de los préstamos, pues indicaba cuales empresas presentan poca capacidad de evaluación y cuales empresas garantizaban una capacidad de evaluación adicional por su posición fuera de la zona de ignorancia.

Respecto a los porcentajes de error de clasificación, obtuvo en promedio un 5% para un año previo a la quiebra (tipo I = 6%; tipo II = 3%); un 17 % para dos años previos (tipo I = 28 %, y tipo II = 6 %). Sin embargo, para el tercero, cuarto y quinto año, los porcentajes de error se incrementaron significativamente, obteniéndose 52 %, 71 % y 14 % respectivamente.

Después de obtener estos resultados, Altman seleccionó una segunda submuestra de validación compuesta por 25 empresas en quiebra, las cuales dieron como resultado un porcentaje de error general de sólo el 4% un año previo a la quiebra. Para otra submuestra de 66 empresas fracasadas, aunque esta vez estas no estaban en quiebra, sino sólo con agudos problemas financieros, se obtuvo un porcentaje de error del 21%, cinco años previos a la quiebra.

Volviendo a la muestra inicial, en donde obtuvo el 95 % de aciertos un año previo a la quiebra, Según Altman, su modelo predictivo proporcionaba un pronóstico más exacto hasta dos años previos a la quiebra, pero conforme pasaban los años la exactitud tendía a perderse.

Es muy importante hacer notar que para Altman la cuestión principal no radicaba en si la empresa tenía que acabar forzosamente fracasando por tener unos niveles determinados en sus ratios, sino la importancia de su modelo radicaba en si sus síntomas eran similares a los de otras empresas que sí avanzaban o terminaban en un proceso de fracaso. Es decir, el modelo predictivo de Altman era de alerta y previsión.

En general, los críticos del modelo Altman señalaban que las variables independientes del modelo inicial no eran las más representativas. De ahí que posteriores estudios al seleccionar las variables independientes se apoyasen más en el trabajo de Beaver, y únicamente para el aspecto metodológico se tomaba como referencia a Altman.

Por ejemplo, la variable X5 propuesta por Altman y que se refiere al factor de rotación o eficiencia global de activos (ventas / activo total) fue considerado como un ratio no significativo en los modelos univariados, pues apenas tenía capacidad de discriminación. En cambio, en los modelos multivariados representaba la segunda variable independiente más



importante de la función. Según Joy Tolleferson, esto se debía a un defecto metodológico en la medición correspondiente a las contribuciones individuales que son deducidas mediante los valores de los coeficientes estandarizados.

También se cuestionaba mucho el problema que presentaba la variable X4 en aquellas empresas que no cotizan, pues varios analistas opinaban que al no cotizar en Bolsa la mayoría de las firmas, la función era poco práctica al requerir indicadores del mercado de valores.

Con respecto a la X4, Lizarraga (1993) opinó que existían dos importantes situaciones que fueron: a) El ratio X2 (beneficios retenidos / activo total), que recogía el tipo de información de la X4; y b) el ratio X4 que no representaba un verdadero indicador del endeudamiento, a excepción de aquellas empresas que tuvieran una gran dependencia del mercado de valores y dada su volátil cotización podría desvirtuar el sentido del ratio.

Algunos estudios como los de Moyer (1977) habían logrado en la reestimación eliminar a las variables X4 y X5.

Altman propuso entonces cambiar a datos contables el numerador de la variable X4 para aquellas empresas que no cotizaran en la bolsa y revisar la variable X5 para su posible eliminación del modelo. A partir de las observaciones de Jhonson (1970) y Moller (1977), Altman llegó a comprender algunas limitaciones y defectos en su modelo; sobre todo en lo tocante a lo polémica sobre la capacidad predictiva.

Conclusiones - La precisión de la Altman Z-score en la predicción de quiebras

La precisión de la fórmula Altman Z-score es de un 72% con 2 años de antelación con respecto a la fecha de la quiebra, con un porcentaje de falsos negativos del 6%.

En un período de prueba de 31 años, esta fórmula tuvo una precisión de entre un 80% y un 90% a la hora de predecir quiebras un año antes de que sucediesen, con un porcentaje de falsos negativos de entre un 15% y un 20%.

Por lo tanto, podemos decir que la fórmula de Altman para predecir quiebras es bastante precisa. No obstante, no es una fórmula infalible, por lo que debemos combinar su utilización con un análisis cualitativo de la empresa que ayude a determinar si estamos ante una futura empresa en quiebra.¹⁶

Algunas desventajas al modelos Altman

Otras desventajas del modelo, según Jhonson (1970), se referían a la poca capacidad de los ratios financieros para llevar predicciones "ex ante". En cambio, cuando las predicciones eran "ex post", el modelo Altman tendía a ser más exacto y se reconocían más las causas del fracaso a través de los estados financieros. Sin embargo, el modelo no por eso podía diferenciar realmente a las empresas que iban a fracasar, ya que no todas las clasificadas como fracasadas llegaban a su fin.

El modelo predictivo también tenía poca capacidad para captar el dinamismo del proceso del fracaso empresarial, y los ratios por sí solos no tenían la capacidad para describir el proceso

¹⁶ Francisco Javier Lodeiro Amador Analista bursátil - <http://www.academiadeinversion.com/altman-z-score-formula-predecir-quiebras/>



dinámico de la quiebra, pues éste hecho hasta la fecha es muy complejo en términos cuantitativos y cualitativos. Tradicionalmente los ratios sólo habían servido para los análisis comparativos estáticos. El mismo Altman reconoció el problema del dinamismo en estos modelos.

En general, los modelos como los de Altman, aunque habían demostrado que las empresas fracasadas y las sanas presentaban ratios diferentes, no habían demostrado estos concluyentemente que tuvieran poder predictivo.

En cuanto a la variable dependiente, existía la polémica sobre la necesidad de delimitar la definición de fracaso, pues si sólo se incluía un término, esto producía continuos errores de estimación en el modelo. Por otra parte, si la definición se ampliaba a varios términos, entonces el modelo era más exacto estadísticamente pero más subjetivo para el mundo real.

Los modelos basados en ratios como el de Altman, y en el caso de economías emergentes, presentan problemas fundamentales en cuanto a la calidad y disponibilidad para obtener bases de datos fiables.¹⁷

Aproximación tecnológica - procesos de desarrollo

Tal como se mencionó en la parte introductoria del trabajo, uno de los objetivos del proyecto, se basa en evaluar la eficacia de los modelos analizados, aplicando datos financieros contemporáneos.

A continuación se realiza una aproximación descriptiva sobre aspectos tecnológicos que consideramos relevantes para el desarrollo de esta actividad.

XBRL

Tal como se mencionó anteriormente, la SEC¹⁸ de USA ha puesto a disposición un amplio repositorio con los Estados Contables de las empresas americanas que regula. Esta información se encuentra codificada con el lenguaje XBRL (eXtensible Business Reporting Language)¹⁹. XBRL es un lenguaje derivado de XML (eXtensible Mark-Up Language), estándar universalmente aceptado para la representación, transmisión y explotación de información en Internet, reglado y promovido por la w3c²⁰.

El marco de trabajo (Framework) que provee XBRL para la representación de información financiera y de negocios se sustenta elementalmente en tres ámbitos de trabajo: documentos de instancias (reportes XBRL), Taxonomías y documentos fundacionales.

Los documentos fundacionales son los esquemas de datos elementales para la codificación de información financiera. En estos documentos se definen los tipos de datos que utilizará el

¹⁷ <http://www.academiadeinversion.com/altman-z-score-formula-predecir-quiebras/>

¹⁸ SEC: Securities and Exchange Commission (2016) - "SEC-XBRL" version obtenida el 26/07/2016 <https://www.sec.gov/xbrl/site/xbrl.shtml>

¹⁹ XII XBRL International Inc.(2013) Especificación XBRL 2.1 versión obtenida el 26/07/2016 <http://www.xbrl.org/Specification/XBRL-2.1/REC-2003-12-31/XBRL-2.1-REC-2003-12-31+corrected-errata-2013-02-20.html>

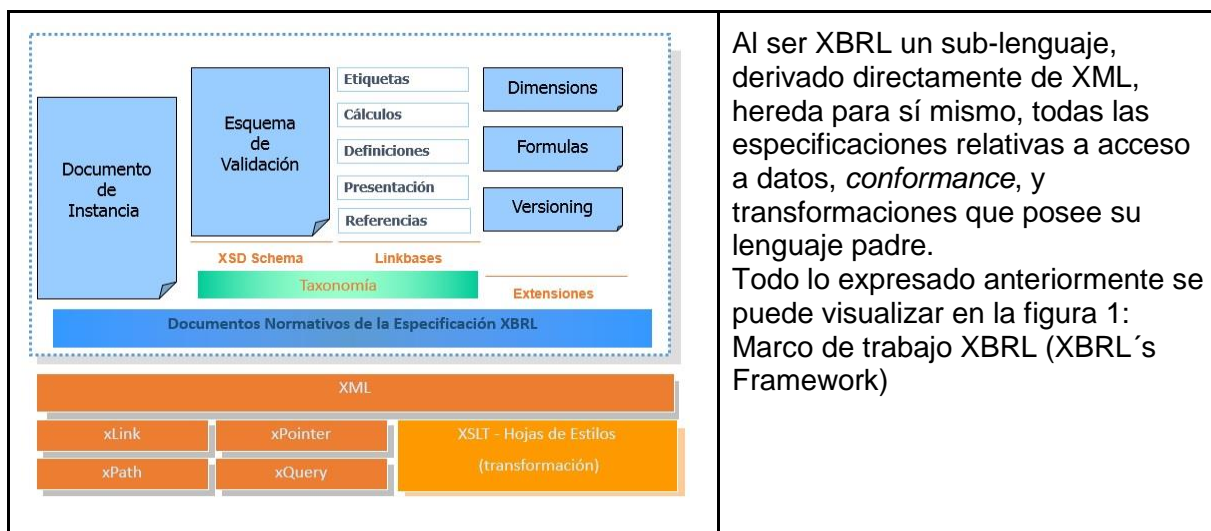
²⁰ W3C World Wide Web Consortium Especificación XML version obtenida el 26/07/2016 <https://www.w3.org/TR/2008/REC-xml-20081126/>



lenguaje (datos monetarios, formatos de fechas, etc..)

Las Taxonomías de XBRL se encuentran compuestas por 2 tipos de recursos. Los esquemas de datos (XML Schemas), en los cuales se definen los elementos financieros a ser utilizados, y los "linkbases" que estructuran las relaciones que existen entre esos elementos financieros (relaciones de cálculos, de definiciones, de orden de presentación, etc.).

Figura 2: Marco de Trabajo XBRL (XBRL's Framework)



Por último, los documentos de instancias o reportes XBRL, son los que contendrán la información financiera particular de un organismo a un momento determinado. Esta información será validada contra la Taxonomía para asegurar la consistencia de la información financiera representada en los mismos. Es de destacar que el marco de trabajo de XBRL también provee especificaciones para el modelado dimensional y multidimensional de datos financieros (XBRL dimensions)²¹, versionado taxonómico (versioning)²², procesos de rendering (XBRL InLine)²³ y desarrollos para validaciones de cálculos y relaciones complejas de elementos financieros (XBRL fórmulas)²⁴.

JSON

²¹ XII XBRL International Inc.(2012) XBRL Dimensions 1.0 versión obtenida el 26/07/2016 <http://www.xbrl.org/specification/dimensions/rec-2012-01-25/dimensions-rec-2006-09-18+corrected-errata-2012-01-25-clean.html>

²² <http://www.xbrl.org/specification/versioning-base/rec-2013-02-27/versioning-base-rec-2013-02-27.html> XII XBRL International Inc.(2013) Especificación XBRL Versioning 1.0 versión obtenida el 26/07/2016

²³ XII XBRL International Inc.(2013) Especificación XBRL InLine 1.1 versión obtenida el 26/07/2016 <http://www.xbrl.org/specification/inlinexbrl-part1/rec-2013-11-18/inlinexbrl-part1-rec-2013-11-18.html>

²⁴ XII XBRL International Inc.(2013) Especificación XBRL Formula 1.0 versión obtenida el 26/07/2016 <http://www.xbrl.org/wgn/xbrl-formula-overview/pwd-2011-12-21/xbrl-formula-overview-wgn-pwd-2011-12-21.html>



En julio de 2006, Douglas Crockford publicó la especificación oficial de formato del lenguaje JSON (JavaScript Object Notation)²⁵. El lenguaje JSON fue ganando amplia aceptación en el ámbito de tecnología informática ya que es una versión más simplificada e intuitiva (human readable) de XML. Originariamente la motivación de su desarrollo se centró en la simplificación de interoperabilidad entre servidor y cliente en páginas web. Al ser un lenguaje basado en la notación de JavaScript, actúa con facilidad en los procesos vinculados a componentes embebidos en páginas web.

En la actualidad, si bien se cuenta con bases de datos nativas en XML (por ejemplo eXist²⁶), la mayoría de las bases de datos NoSQL de tipo documentales, sobre las que nos referiremos más adelante, utilizan como estándar de codificación de datos el lenguaje JSON.

La transformación de reportes y esquemas de XBRL en documentos con formato JSON, es una de las actividades relevantes en el proceso de desarrollo del proyecto SECBI.

Con esa finalidad se prevé en el proyecto SECBI realizar la conversión de formatos XBRL a JSON utilizando los lineamientos de la especificación (en proceso) "xBRL-JSON: mapping from Open Information Model 1.0" de XBRL International Inc²⁷.

Big Data

El crecimiento exponencial de los datos que se gestionan en la Web, en especial motivado por la irrupción de redes sociales, ha derivado en la generación de un nuevo paradigma de administración de datos, al que se le denomina comúnmente "Big Data".

La mayoría de los autores coinciden en delinear 3 ejes directrices en el concepto de Big Data (las tres V):

Volumen: Cantidades masivas de datos, generalmente medidas en Gb, Tb, Pb.

Velocidad: Hace referencia a la disponibilidad de esa cantidad masiva de datos para poder ejecutar sobre los mismos búsquedas o procesos de análisis.

Variedad: Al tener primacía en la conformación de estos grandes repositorios de datos, información derivada de redes sociales, se debe contemplar la incorporación de datos estructurados, semi-estructurados y no estructurados.

²⁵ JSON, versión obtenida el 27/07/2016 <http://www.json.org/>

²⁶ EXIST DB, versión obtenida el 27/07/2016 <http://exist-db.org/exist/apps/homepage/index.html>

²⁷ XBRL-JSON: mapping from Open Information Model 1.0 - XBRL International Inc., versión obtenida el 27/07/2016 <http://www.xbrl.org/Specification/xbrl-json/PWD-2016-01-13/xbrl-json-PWD-2016-01-13.html>



Figura 3: Tres ejes directrices de Big Data.



Nuestro enfoque en este punto se centrará en dos ejes que consideramos relevantes para el desarrollo del proyecto SECBI: Volumen y variedad de datos.

Grandes volúmenes de datos: Actualmente el repositorio de datos disponibilizado por la SEC cuenta con 187.000 presentaciones, que consumen 704 Gb de almacenamiento en disco. Una métrica estimada de este volumen de información en una base de datos relacional ronda entre 40 a 50 millones de registros, lo cual conlleva una significativa caída de performance de funcionamiento, salvo que se utilice una Base de Datos relacionales (RDBMS) de alta escalabilidad empresarial.

Variedad de datos: En referencia al otro eje directriz de Big Data que destacamos, el de variedad de datos, es importante destacar que si bien la información financiera de empresas que facilita la SEC, se encuentra codificada con el estándar XBRL, el sistema EDGAR (Electronic Data Gathering, Analysis, and Retrieval System) posibilita a las empresas reguladas reportar información semi estructurada por medio del mecanismo de extensibilidad taxonómica de XBRL.

La SEC en concordancia con la FASB (Financial Accounting Standards Board²⁸) de USA pone a disposición de sus empresas reguladas la Taxonomía XBRL que contiene el esquema de datos que deben utilizar para reportar sus Estados Contables. Sin embargo este sistema no se restringe a un esquema cerrado, sino que permite a las empresas “heredar” la taxonomía FASB (US-GAAP), y “extender” la misma por medio de módulos taxonómicos desarrollados por las empresas reguladas, en los cuales pueden agregar cuadros de información financiera o brindar un mayor detallamiento o desagregación de la información contemplada por la Taxonomía base.

Modelo relacional alternativo

Sin desmedro de la elección realizada para adoptar el sistema de BD NoSQL MongoDB que se explicó anteriormente, se ha considerado incorporar dentro del desarrollo del proyecto

²⁸FASB Financial Accounting Standards Board, versión obtenida el 27/07/2016 <http://www.fasb.org/home>



SECBI, una interfaz de transformación de datos NoSQL (en formato JSON en nuestro caso) a datos en formato relacional. La motivación de implementar esta interfaz es facilitar la implementación de consultas y procesos de análisis de datos.

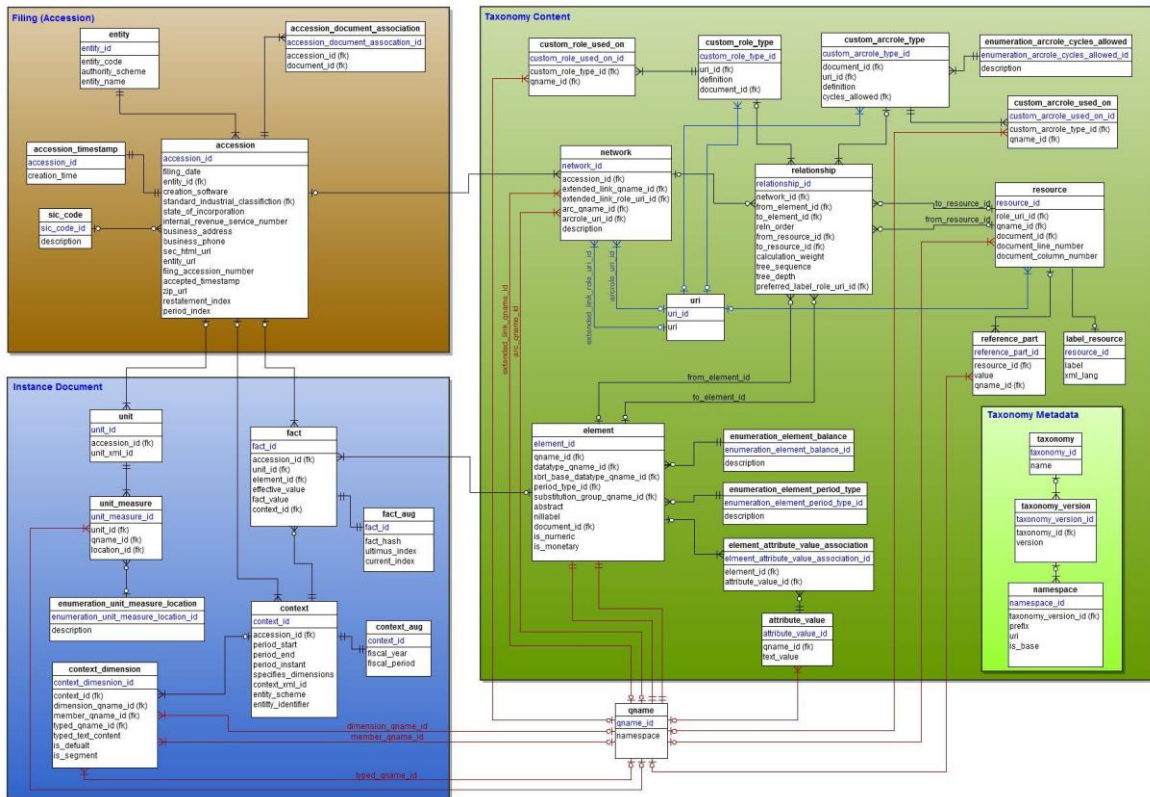
Por su capacidad de procesamiento distribuido MongoDB facilita la gestión masiva de datos. Esto se complementa con una alta versatilidad para gestionar cambios en la estructura de datos, derivada de su condición de BD schema-less. Sin embargo tal como su nombre lo indica, las BD NoSQL no cuentan con esquemas estructurados de datos, y necesitan de implementar métodos de consulta de información diferentes al lenguaje SQL (Structured Query Language) que utilizan las tradicionales bases de datos relaciones.

Tomando en consideración la universalidad de adopción que posee en el ámbito actual el lenguaje SQL, se prevé, dentro del marco del proyecto SECBI la implementación de una interfaz NoSQL-relacional.

Las consultas de tipo "genéricas" se ejecutarán sobre la totalidad del repositorio de datos NoSQL MongoDB, por ejemplo: "todas las empresas con determinado código de actividad", "empresas industriales entre 2012 y 2015", etc...

El resultado de esta consulta conformará un "sub-set" de datos NoSQL que se transformarán a formato relacional, a fin de realizar los procesos de inteligencia de negocios, machine learning, o data mining sobre RDBMS como MySQL, SQL Server, Oracle, etc...

Para desarrollar y vincular esta transformación de datos NoSQL a datos relacionales, tenemos previsto utilizar el modelo abstracto de datos propuesto por la SEC para gestión de datos en RDBMS. Dicho modelo puede ser visto en la Figura 4.





Aplicación de técnicas de análisis, machine learning, data mining y otras

Las últimas actividades que prevé el proyecto a los fines de lograr sus objetivos, se vinculan a la utilización de técnicas de inteligencia de negocio (data management, análisis exploratorio, etc...) y de descubrimiento de conocimiento (machine learning, data mining, etc...).

El objetivo particular de implementar estas técnicas se encuentra vinculado al descubrimiento de relaciones insospechada de datos que puedan generar modelos de evaluación de gestión empresarial, como así también nuevos variables de predictibilidad de insolvencia empresarial.

Algunas de las técnicas más comunes que se espera analizar e implementar se enumeran en el cuadro siguiente:

Procesos y técnicas de análisis a evaluar / implementar

Proceso de Análisis	Descripción	Algoritmos utilizados	Grado de dificultad
Data Management	Proceso previo al análisis de datos, donde se preparan los data sets para ser estudiados	NAs (Tratamiento de valores faltantes) Eliminación de ruido (noise reduction) Normalización Transformación de datos Detección de outliers	Bajo
Análisis Exploratorio	Técnicas estadísticas básicas al que se someten los datos para determinar relaciones, asociaciones, agrupamientos básicos y así obtener conclusiones preliminares en el análisis	Correlación de variables Intervalos de confianza Histogramas Densidad	Medio
Visualización de datos	Técnicas de comunicación visual de datos, basadas en la habilidad del cerebro humano de interpretar con mayor facilidad información presentada visualmente	Regresión Lineal Regresión local Scatter plots Box plots Cuartiles Gráficas de barras Gráficas de densidad	Medio

///



Aprendizaje de Máquinas (Machine Learning)	Técnicas que permiten a sistemas aprender de datos, por medio de un entrenamiento basado en resultados observados con anterioridad	Regresión Linear Regresión Logística Least squares regression k-nearest Naïve Bayes K-means clustering Árboles de decisión Máquinas de vectores de soporte Random forests	Alto
Data Mining	Proceso no trivial de identificación de patrones válidos, novedosos, potencialmente útiles y comprensibles en los datos (Witten, Frank, Hall, 2011)	Análisis de Cluster: K-means clustering K-medoids clustering Hierarchical clustering Estimaciones de densidad Expectation-maximization Reglas de Asociación Patrones: Eclap Apriori Sequence metrics	Alto
Text Mining	Técnica especializada de Data Mining, enfocada a grandes cantidades de datos, en formato de texto, no estructurado, masivo y de mucha variabilidad	Procesamiento de textos Clustering de textos	Alto

Conclusiones y propuesta

El modelo analizado, surgido de la comparación entre distintos autores²⁹, resulta un buen inicio para obtener las primeras conclusiones y proponer alternativas acordes con nuestra realidad, en lo geográfico y en lo temporal.

Los cambios producidos por la irrupción tecnológica y por la globalización de la economía, han derivado en nuevas técnicas, formas de asociaciones, acceso a mercados financieros y procesos de negocios adoptados por las empresas contemporáneas. Esta realidad presenta

²⁹ Marchese, A.; Díaz D.; Campanaro R.; Carrizo, F.; Ferraro, A.; Scialla J.; Diaz Toledo S.; Rivero J. (2015) - USO DE HERRAMIENTAS DE INTELIGENCIA DE NEGOCIO PARA EL ANÁLISIS DE INFORMES FINANCIEROS - IX Jornadas de Investigación en Ciencia y Tecnología UNR



grandes cambios en referencia a las situaciones analizadas por Altman en los años 60/70, lo cual abre un ámbito de discusión sobre si se deberían rever estos modelos bajo las actuales circunstancias.

El diagnóstico propuesto por Altman, categoriza a las empresas bajo análisis en 3 grupos: "saludables, zona gris y enfermas". Sin embargo, en el caso de que una empresa se incluya dentro de las "saludables", no nos muestra con claridad la eficiencia en la gestión de la misma. Esto surge de que el objetivo buscado por Altman, fue identificar las empresas en riesgo de quiebra o insolvencia. Es nuestra intención continuar esta investigación con un análisis específico de evaluación de gerenciamiento, a través de indicadores de gestión (KPI - Key Performance Indicators).

BIBLIOGRAFIA

Altman Edward (1995) Corporate Credit Scoring Models: Approaches and Standards for Successful Implementation - New York University (NYU) - Salomon Center; New York University (NYU) - Department of Finance

Altman Edward (2002) Revisiting Credit Scoring Models in a Basel 2 Environment New York University (NYU) - Salomon Center; New York University (NYU) - Department of Finance May 2002

Altman Edward I. (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate

Bell, J., 2015. Machine Learning: Hands-On for Developers and Technical Professionals. Willey.

Bergeron, B. (2004). Essentials of XBRL: Financial reporting in the 21st century (Vol. 30). John Wiley & Sons.

Colin Ware, E., 2004. Information Visualization: Perception for Design. Morgan Kaufmann.

Foster Provost, T. F., 2013. Data Science for Business. O'Reilly

Frampton, M. (2014). Big Data Made Easy: A Working Guide to the Complete Hadoop Toolset. Apress.

Harrington, P., 2012. Machine Learning in Action. MANNING.

Hoffman, C., & Watson L. (2009). XBRL for Dummies. John Wiley & Sons.

Ian H. Witten, E. F. M. A. H., 2011. Data Mining: Practical Machine Learning Tools and Techniques. s.l.:Elsevier / Morgan Kaufmann.

Jackson, W. (2016), JSON Quick Syntax Reference, Apress, Lompoc, California, USA.

Aldazábal Contreras, Janet Cecibel -Alberto Fernando Napán Vera Revista de la Facultad



de Ciencias Contables Vol. 22 N. ° 42 pp. 53-59 (2014) UNMSM, Lima - Perú ISSN: 1560-9103 (versión impresa) / ISSN: 1609-8196 (versión electrónica).

Kuncheva, L. I., 2014. Combining Pattern Classifiers. Second Edition ed. s.l.:Wiley.

Marchese, A.; Díaz D.; Campanaro R.; Carrizo, F.; Ferraro, A.; Scialla J.; Diaz Toledo S.; Rivero J. (2015) - USO DE HERRAMIENTAS DE INTELIGENCIA DE NEGOCIO PARA EL ANÁLISIS DE INFORMES FINANCIEROS - IX Jornadas de Investigación en Ciencia y Tecnología UNR

Mohanty, S., Jagadeesh, M., & Srivatsa, H. (2013). Big Data Imperatives Apress.
Morrison, M. (2000). XML al descubierto. Prentice Hall. Madrid.

Ohlson, James A. "Financial Ratios and the Probabilistic of Prediction of Bankruptcy". Publicado en Journal Accounting of Research, Vol.18, No. 1. Spring (1980), pp. 109-131. JTSOR (LIBRE ACCESO).

Ringelin (2004) "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy"

Rischpater, R. (2015). JavaScript JSON Cookbook. Packt Publishing Ltd.

Stubbs, E. (2014). Big Data, Big Innovation: Enabling Competitive Differentiation Through Business Analytics. John Wiley & Sons.

Taffler, R. (1983). " La evaluación de la solvencia y el rendimiento de la empresa mediante un modelo estadístico ". Contabilidad y Business Research, Otoño, pp. 295-307

Tufte, E. R., 1990. Envisioning Information. s.l.:GRAPHICS PRESS.

Van Horen, J. C. W., & John, M. Fundamentos de administración financiera/por James C. Van Horne y John M. Wachowicz (No. 658.15 V3y 1992.).

Yu-Wei, Chiu (David Chiu), 2015. Machine Learning with R Cookbook. s.l.:Pack Publishing.

Ibarra Mares, Alberto - "Una perspectiva sobre la evolución en la utilización de las razones financieras o ratios "pensamiento y gestión, n° 21 universidad del norte issn 1657-6276

Citas de Internet

FASB Financial Accounting Standards Board, versión obtenida el 27/07/2016 <http://www.fasb.org/home>

SEC: Securities and Exchange Commission (2016) - "SEC-XBRL" version obtenida el 26/07/2016 <https://www.sec.gov/xbml/site/xbml.shtml>.

XBRL Versioning 1.0 versión obtenida el 26/07/2016



<http://www.xbrl.org/specification/versioning-base/rec-2013-02-27/versioning-base-rec-2013-02-27.html>

XII XBRL International Inc.(2012) XBRL Dimensions 1.0 versión obtenida el 26/07/2016
<http://www.xbrl.org/specification/dimensions/rec-2012-01-25/dimensions-rec-2006-09-18+corrected-errata-2012-01-25-clean.html>

XII XBRL International Inc.(2013) Especificación XBRL 2.1 versión obtenida el 26/07/2016
<http://www.xbrl.org/Specification/XBRL-2.1/REC-2003-12-31/XBRL-2.1-REC-2003-12-31+corrected-errata-2013-02-20.html>.

XII XBRL International Inc.(2013) Especificación XBRL Formula 1.0 versión obtenida el 26/07/2016
<http://www.xbrl.org/wgn/xbrl-formula-overview/pwd-2011-12-21/xbrl-formula-overview-wgn-pwd-2011-12-21.html>

XII XBRL International Inc.(2013) Especificación XBRL InLine 1.1 versión obtenida el 26/07/2016
<http://www.xbrl.org/specification/inlinexbrl-part1/rec-2013-11-18/inlinexbrl-part1-rec-2013-11-18.html>

Ibarra Mares, Alberto - ANÁLISIS DE LAS DIFICULTADES FINANCIERAS DE LAS EMPRESAS EN UNA ECONOMÍA EMERGENTE: LAS BASES DE DATOS Y LAS VARIABLES INDEPENDIENTES EN EL SECTOR HOTELERO DE LA BOLSA MEXICANA DE VALORES -<http://www.eumed.net/tesis-doctorales/2010/aim/MODELO%20ALTMAN.htm> - obtenida en setiembre 2016

Garcia Gallego Ana, Mures Quintana Maria Jesús "La muestra de empresas en los modelos de predicción del fracaso: influencia en los resultados de clasificación" Revista de métodos cuantitativos para la economía y la empresa Pág. 135 Junio 2013 ISSN: 18869-516X-<http://www.upo.es/RevMetCuant/pdf/vol15/art73.pdf> - obtenida en setiembre 2016

León Valdés, César A. - "El análisis financiero como herramienta en la predicción de quiebra e insolvencia financiera.Universidad Externado de Colombia
<http://revistas.uexternado.edu.co/index.php/contad/article/view/1295/1232> - obtenida en setiembre 2016

<http://www.academiadeinversion.com/altman-z-score-formula-predecir-quiebras/> - obtenida en setiembre 2016