

Modelos de predicción de quiebra financiera

1

Juan Tovar*, David Mautino*, Rosmery Aguilar*, Luis Enciso*

Resumen—La evaluación de diferentes tipos de indicadores financieros para la predicción de quiebra en compañías es un tema de investigación que viene ganando importancia debido a la mayor disponibilidad de datos abiertos; más aún se ha demostrado en trabajos anteriores la alta eficacia de algunos métodos de aprendizaje máquina para la estimación de dicha condición. En el presente trabajo se busca identificar los modelos de aprendizaje máquina más óptimos para la realización de predicción de quiebra financiera basado en indicadores económicos que puedan ser en adelante adecuados a la realidad financiera peruana.

Index Terms—quiebra financiera, aprendizaje máquina, indicadores financieros

I. INTRODUCCIÓN

LA estimación efectiva de bancarrota en empresas es vital en muchos aspectos de la vida económica nacional, e.g. en muchas instituciones financieras se considera para decisiones de préstamo; además la bancarrota económica tienen un efecto no solo en la misma empresa, sino también en la economía global. Por los efectos e intereses mencionados existe una gran necesidad de generar modelos de predicción de bancarrota financiera; además la disponibilidad de abundante data financiera histórica permite que la generación de modelos predictivos pueda realizarse basado en modelos de aprendizaje máquina, lo que es reforzado por el buen desempeño mostrado por tales técnicas en tareas similares.

El objetivo de este estudio es por ello la generación de modelos de aprendizaje máquina adecuados para la predicción de quiebra financieras en empresas, bajo las siguientes condiciones: las empresas deben tener un mínimo de tres años de información financiera disponible, y debe existir un número suficiente de empresas en la misma área industrial para la comparación de casos de quiebra y no quiebra. Como objetivos específicos se consideran, la realización de una selección adecuada de indicadores económicos para el entrenamiento, la generación de modelos y el análisis comparativo de sus desempeños; y la la evaluación de modelos de ensamble.

El presente informe se compone de las siguientes secciones; en la sección 2 se describe el estado del arte en la generación de modelos de predicción de quiebra basado en indicadores financieros, y en la sección 3 se describe el diseño del experimento, mencionando las consideraciones de selección de indicadores financieros y se describe el data-set, el

preprocesamiento e imputación del data-set, el procedimiento para el entrenamiento y los criterios de evaluación a utilizar.

II. ESTADO DEL ARTE

Las primeras aproximaciones a la predicción de quiebra financiera se dieron basadas en estudios empírico que permitían determinar el nivel de solvencia con el fin de predecir o evitar la posible quiebra a través de utilización de ratios, métodos estadísticos y el análisis financiero sobre los componentes de la solvencia, liquidez, rentabilidad, entre otros factores, numerosos estudios a partir de 1932 [1] han elaborado este aproximamiento. En los últimos años, la resurgencia de los métodos de aprendizaje máquina ha generado un enorme interés en la generación de modelos basados en aprendizaje supervisado. En [2] Liang presenta el data-set de indicadores financieros (FRs) basado en tres años de actividad empresarial en Taiwan, con el objetivo de utilizarlos en conjunto con indicadores de gobierno corporativo (CGI), no obstante, estos últimos son específicos a la realidad de ese país. Muchos trabajos en la materia han demostrado que la utilización exclusiva de FRs puede alcanzar una precisión alta, en [3] se presenta un resumen de los modelos aplicados a esta tarea de predicción usando indicadores financieros usando métodos estadísticos y aprendizaje máquina, e indica que estos últimos reportan mejor rendimiento; además se brinda una mirada comparativa a la aplicación de Redes Neuronales Artificiales (ANN), Regresión Logística (LR), Árboles de decisión (DT), y Matrices Vectoriales de Soporte (SVM), en la comparación SVM logra mejores resultados, además se puede realizar una optimización de hiper-parámetros mediante algoritmos evolutivos como Algoritmos Genéticos (GA) y Optimización por Enjambre de Partículas (PSO). En muchas ocasiones se realiza una mejora basada en modelos conjuntos (ensemble), usando técnicas tales como bagging, boosting, stacking, y sistemas de votaciones (e.g. mayoritaria, unánime). En [4] Liang propone un esquema de ensamble basado en voto unánime (UV) aplicado al data-sets como el de [2], se reporta el mejor modelo combinado de SVM, ANN basado en bagging y CART basado en boosting, este enfoque es mejorado en [5] con un esquema de stacking ensemble para integrar un modelo basado en FRs con otro modelo basado en CGIs y score-Z, definido en [6], en este último trabajo [5] se usa Discriminante Analítico de Pasos (SDA) para la selección de FRs y CGIs y SVM para los modelos. Otros trabajos recientes tales como [7], donde Lin presenta un análisis para tres data-sets, muestra que ningún modelo es superior de manera absoluta en los tres casos, no obstante se reporta que los mejores modelos obtenidos son Naive Bayes (NB) + GA y SVM + GA, además ellos pueden complementarse con bagging; mientras que en [8] Tang

¹INF648

*Pontificia Universidad Católica del Perú, Escuela de Posgrado, Maestría de Informática (e-mail: author@pucp.edu.pe).

Aprendizaje Profundo: Teoría y aplicaciones (INF648)

presenta una comparativa de modelos de predicción basado en FRs y factores textuales y de manejo empresarial, estos factores son separados en sub-grupos dependiendo de su naturaleza y los modelos obtenidos son comparados, se puede observar una ligera mejora al incluir estos factores, sobretodo en modelos de ensemble por stacking y de Redes Neuronales Recurrentes (RNN).

III. DISEÑO DEL EXPERIMENTO

El diseño del experimento se realizó tomando en consideración variables financieras de la empresas de Taiwán obtenidas del Taiwan Economic Journal² en los años 1999 a 2009 que nos indicará cual ha sido el comportamiento de las empresas y además nos indica cuando una empresa quebró. La quiebra de la empresa está definida en base a las regulaciones comerciales de la Bolsa de Valores de Taiwán.

3.1 Descripción del conjunto de Datos

El dataset consta de 6819 casos y 96 variables, se valida que no se tienen datos faltantes en cada variable.

El dataset del experimento será normalizado en el rango de 0 a 1, en base a la siguiente fórmula:

$$\forall x \in F, \text{normalize}(x) = \frac{x - \min(F)}{\max(F) - \min(F)},$$

donde F es un conjunto de una característica específica (es decir, variable), x es el valor de la característica y max (F) y min (F) son los valores máximo y mínimo del conjunto de características específicas, luego se observa que la distribución de tipo de característica en las variables financieras del dataset es el siguiente:

- 3 del tipo binario, que incluye al dato indicador de quiebra (flag)
- 93 del tipo continuo y
- Ninguno de tipo discreto.

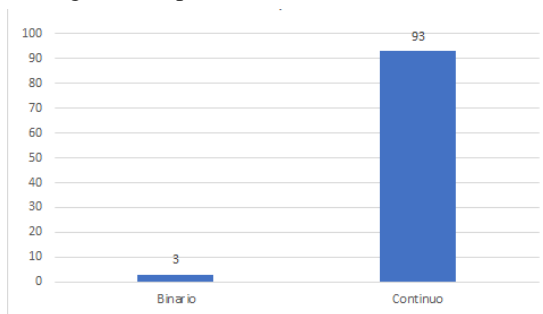


Figura 1. Dataset Taiwanes Tipo de Característica en Variables

Debido a que el dato a evaluar (flag) es binario, los modelos serán de clasificación, considerando ello se utilizara una muestra aleatoria del 10% para el conjunto de prueba y el 90% restante para el conjunto de entrenamiento. Se debe resaltar al

igual que el estudio realizado en [2] que existe una diferencia significativa entre el número de casos en quiebra y no en quiebra:

Tabla 1. Diferencia casos en quiebra y no en quiebra

Flag	descripción	Cantidad	Porcentaje
0	No en quiebra	6599	97%
1	En quiebra	220	3%

Esto da como resultado un problema de desequilibrio de clases, que probablemente lleve a una degradación en el desempeño de la predicción final. Por lo tanto, utilizaremos el método de muestreo estratificado [6] para recopilar el mismo número de casos en quiebra y no en quiebra.

3.2 Metodología

3.2.1 Selección de características

El objetivo de la selección de características es determinar las características más representativas, reduciendo las características redundantes o irrelevantes para el apoyo en la optimización del modelo predictivo. Se utilizó el siguiente método para la selección de características:

- Random Forest Classifier

3.2.2 Modelos predictivos

Al tratarse de la predicción de quiebra o no quiebra, se emplearon, en la etapa de experimentación, los siguientes algoritmos de clasificación:

- Decision Tree
- Logistic Regression
- Random Forest Classifier
- XGBoost Classifier
- Genetic Algorithm + Support Vector Machine

3.2.3 Métricas de evaluación

Se utilizaron las siguientes métricas para la evaluación:

- Accuracy
- F1-Score
- Área bajo de la curva ROC
- Type I - error

3.2.4 Optimización de modelos

Se utilizaron los siguientes métodos de optimización:

- Grid Search Cross Validation
- K-Fold Validation

3.2.5 Marco de la metodología

Este estudio se desarrolló en el marco que se muestra en la Figura 2, que implica seis fases claves [11], de las cuales 5

² <https://archive.ics.uci.edu/ml/datasets/Taiwanese+Bankruptcy+Prediction>

Aprendizaje Profundo: Teoría y aplicaciones (INF648)

fueron realizadas en el presente trabajo: comprensión de los datos, preparación de los datos, modelamiento, evaluación y despliegue.

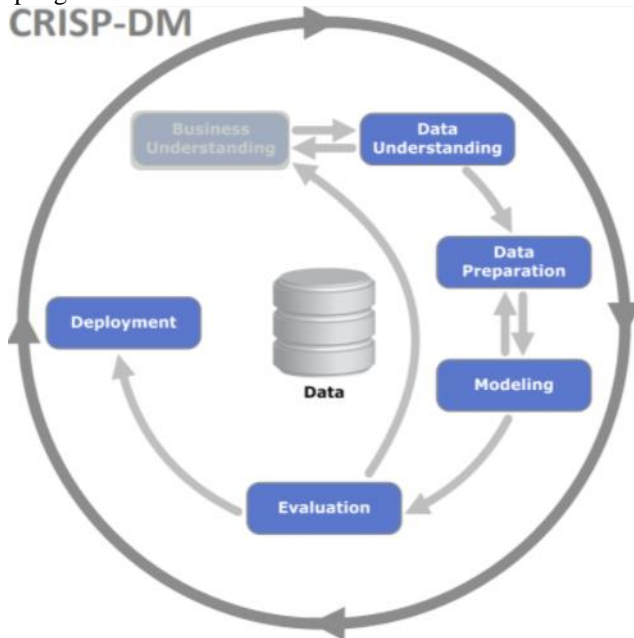


Figura 2. Metodología del desarrollo del modelo a utilizar

IV. EXPERIMENTACIÓN Y RESULTADOS

Las pruebas se realizaron con cada uno de los modelos base, además se realizó el ensamble de los mejores y por último, la búsqueda de un modelo en base a la reducción de las variables más importantes seleccionadas mediante 'feature importance' basado en Random Forest.

En cada caso se usó una división del dataset en sets de entrenamiento y prueba, con 80% y 20% respectivamente, usando el comando `train_test_split` de scikit-learn, en modo estratificado.

4.1 Modelos Base

Se ensayaron los siguientes modelos base:

4.1.1 Logistic Regression (Figura 3)

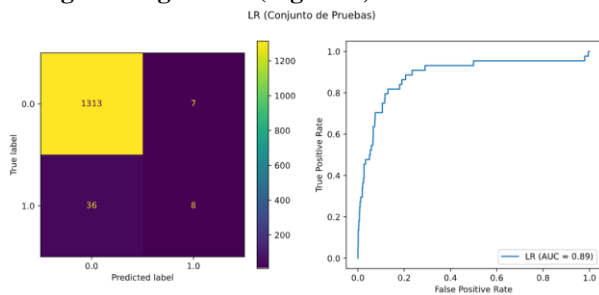


Figura 3. Resultados del modelo de Regresión Logística

4.1.2 Random Forest (Figura 4)

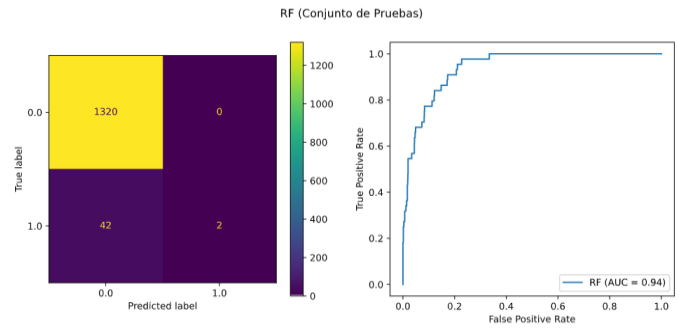


Figura 4. Resultados del modelo de Random Forest

4.1.3 XG Boost (Figura 5)

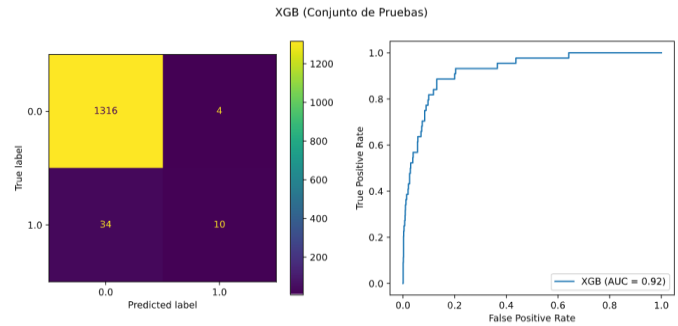


Figura 5. Resultados del modelo XG Boost

4.1.4 Support Vector Machine (Figura 6)

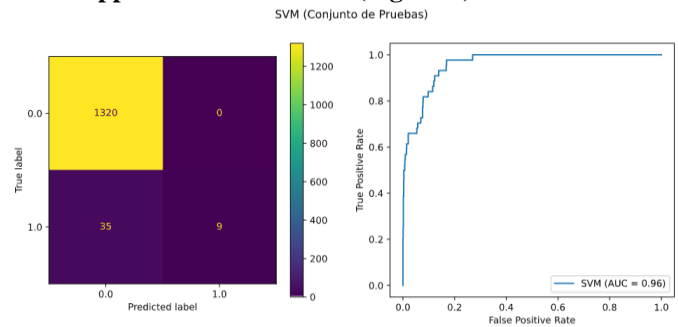


Figura 6. Resultados del modelo SVM

4.1.5 Support Vector Machine + Genetic Algorithm (Figura 7)

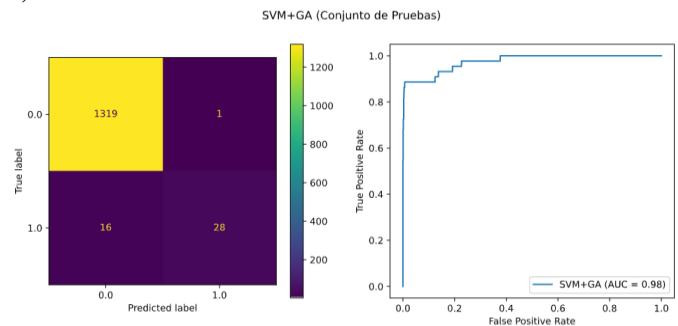


Figura 7. Resultados del modelo SVM + GA

4.2 Modelo Ensamble (Figura 8)

Para el caso de ensamble se ensayó la obtención de un modelo mediante Stacking con optimización por medio de Logistic

Aprendizaje Profundo: Teoría y aplicaciones (INF648)

Regression.

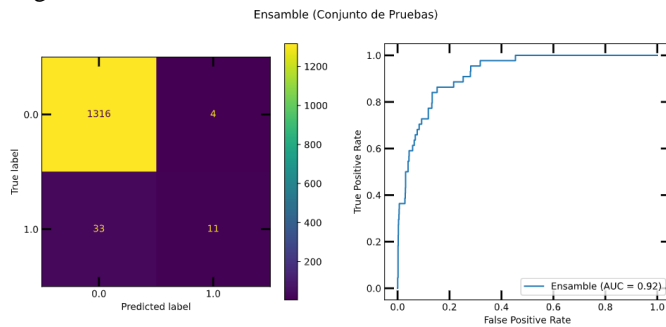


Figura 8. Resultados del modelo ensemble

4.3 Modelos de Implementación

Finalmente se entrenaron modelos con un subgrupo de las características, se encontró que con 70 características se pueden obtener resultados casi tan buenos como el mejor de los modelos previamente entrenados. Para este caso se ensayó con el algoritmo SVM+GA y auto-sklearn [10].

4.1.1 SVM+GA (Figura 9)

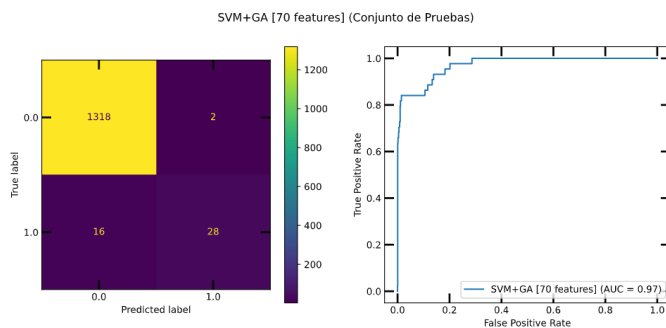


Figura 9. Resultados del modelo SVM+GA

4.1.1 Auto-sklearn (Figura 10)

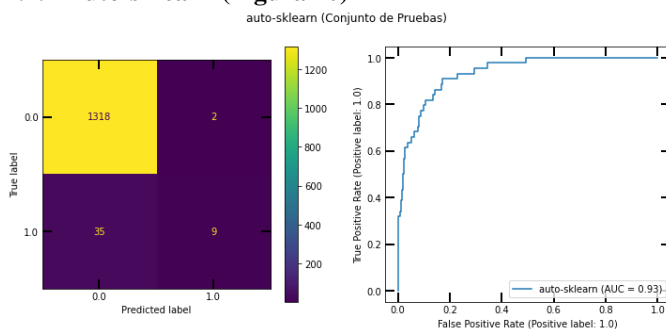


Figura 10. Resultados del modelo Auto-sklearn

V. DISCUSIÓN

- Se observa que los mejores resultados se obtuvieron con el modelo de SVM+GA, con una exactitud del 98.75% y con un F1-score del 76.71% (Tabla 2).

	Pruebas				
	Exactitud	Exhaustividad	F1	AUC	Typel
Ensamble	0.972874	0.250000	0.372881	0.922159	4
SVM	0.974340	0.204545	0.339623	0.958953	0
SVM+GA	0.987537	0.636364	0.767123	0.975207	1
DT	0.968475	0.181818	0.271186	0.885158	7
LR	0.968475	0.181818	0.271186	0.885158	7
RF	0.969208	0.045455	0.086957	0.940771	0
XGB	0.972141	0.227273	0.344828	0.923192	4

Tabla 2. Tabla comparativa de métricas usando las 94 características

- El experimento proporciona una diferencia del modelo auto-sklearn con un F1- score de 32.72% a comparación del modelo SVM+GA con un F1- score de 75.67% . Además se evidencia del despliegue del modelo auto-sklearn con un AUC de 93.35% a comparación del modelo SVM+GA con un AUC de 97.14% (Tabla 3).

	Pruebas			
	Exactitud	F1	AUC	Typel
auto-sklearn	0.972874	0.327273	0.933574	2
SVM+GA [70 features]	0.986804	0.756757	0.971402	2

Tabla 3. Tabla comparativa de métricas usando las 70 características

- Estos resultados deben tenerse en cuenta al considerar 70 features en los modelos SVM-GA y auto-sklearn.

VI. CONCLUSIONES Y TRABAJOS FUTUROS

- Se obtuvieron mejores resultados con el modelo SVM-GA, con una exactitud de 98.7% y una exhaustividad de 63.6%.
- Se generó un modelo ensemble (stacking) con los 4 mejores modelos, sin embargo las métricas no mejoraron.
- Se realizó el feature importance con el algoritmo de Random Forest Classifier y se seleccionaron las 70 más relevantes en el modelamiento de SVM-GA y auto-sklearn.
- El modelo auto-sklearn no mostró mejoras en las métricas, por lo que se sugiere aumentar el tiempo de optimización de la selección del mejor modelo.
- Aplicar métodos más avanzados para solucionar el tema del balanceo del target u obtener mayor cantidad de casos en quiebra.

REFERENCIAS

- [1] FitzPatrick, Paul J., Ph.D. 1932. "A Comparison of the Ratios of Successful Industrial Enterprises With Those of Failed Companies". The Certified Public Accountant Beaver 1968. *Journal of Accounting Research*. (In three issues: October, 1932, p. 598-605; November, 1932, p. 656-662; December, 1932, p. 727-731.
- [2] D. Liang, C. C. Lu, C. F. Tsai y G. A. Shih, «Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study,» *European Journal of Operational Research*, vol. 252, p. 561–572, 2016.
- [3] S. S. Devi y Y. Radhika, «A survey on machine learning and statistical techniques in bankruptcy prediction,» *International Journal of Machine Learning and Computing*, vol. 8, p. 133–139, 2018.
- [4] Liang, D., Tsai, C.F., Dai, A.J. et al. A novel classifier ensemble approach for financial distress prediction. *Knowl Inf Syst* 54, 437–462 (2018). <https://doi.org/10.1007/s10115-017-1061-1>
- [5] D. Liang, C. F. Tsai, H. Y. (Lu y L. S. Chang, «Combining corporate governance indicators with stacking ensembles for financial distress prediction,» *Journal of Business Research*, vol. 120, p. 137–146, 2020.
- [6] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23, 589–609.
- [7] W. C. Lin, Y. H. Lu y C. F. Tsai, «Feature selection in single and ensemble learning-based bankruptcy prediction models,» *Expert Systems*, vol. 36, p. 1–8, 2019.
- [8] Tang X, Li S, Tan M, Shi W. Incorporating textual and management factors into financial distress prediction: A comparative study of machine learning methods. *Journal of Forecasting*. 2020;39:769–787. <https://doi.org/10.1002/for.2661>
- [10] Efficient and Robust Automated Machine Learning, Feurer et al., *Advances in Neural Information Processing Systems* 28 (NIPS 2015).
- [11] R. Wirth and J. Hipp, "CRISP-DM: Towards a standard process model for data mining", *Proc. 4th Intl. Conference on Practical Applications of Knowledge Discovery and Data mining*, pp. 29-39, 2000

Aprendizaje Profundo: Teoría y aplicaciones (INF648)

ANEXO

Dataset

ID	Variable
y	flag
x1	Cost_of_Interest-bearing_Debt
x2	Cash_Reinvestment_Ratio
x3	Current_Ratio
x4	Acid_Test
x5	Interest_Expenses/Total_Revenue
x6	Total_Liability/Equity_Ratio
x7	Liability/Total_Assets
x8	Interest-bearing_Debt/Equity
x9	Contingent_Liability/Equity
x10	Operating_Income/Capital
x11	Pretax_Income/Capital
x12	Working_Capital_to_Total_Assets
x13	Quick_Assets/Total_assets
x14	Current_Assets/Total_Assets
x15	Cash/Total_Assets
x16	Quick_Assets/Current_Liability
x17	Cash/Current_Liability
x18	Current_Liability_to_Assets
x19	Operating_Funds_to_Liability
x20	Inventory/Working_Capital
x21	Inventory/Current_Liability
x22	Current_Liabilities/Liability
x23	Working_Capital/Equity
x24	Current_Liabilities/Equity
x25	Long-term_Liability_to_Current_Assets
x26	Current_Liability_to_Current_Assets
x27	One_if_Total_Liability_exceeds_Total_Assets;
x28	Equity_to_Liability
x29	Equity/Total_Assets
x30	(Long-term_Liability+Equity)/Fixed_Assets
x31	Fixed_Assets_to_Assets
x32	Current_Liability_to_Liability
x33	Current_Liability_to_Equity
x34	Equity_to_Long-term_Liability
x35	Liability_to_Equity
x36	Degree_of_Financial_Leverage
x37	Interest_Coverage_Ratio
x38	Operating_Expenses/Net_Sales
x39	(Research_and_Development_Expenses)/Net_Sales
x40	Effective_Tax_Rate
x41	Book_Value_Per_Share(B)
x42	Book_Value_Per_Share(A)
x43	Book_Value_Per_Share(C)
x44	Cash_Flow_Per_Share
x45	Sales_Per_Share
x46	Operating_Income_Per_Share
x47	Sales_Per_Employee
x48	Operation_Income_Per_Employee
x49	Fixed_Assets_Per_Employee
x50	total_assets_to_GNP_price
x51	Return_On_Total_Assets(C)
x52	Return_On_Total_Assets(A)
x53	Return_On_Total_Assets(B)
x54	Gross_Profit_/Net_Sales
x55	Realized_Gross_Profit/Net_Sales
x56	Operating_Income_/Net_Sales
x57	Pre-Tax_Income/Net_Sales
x58	Net_Income/Net_Sales
x59	Net_Non-operating_Income_Ratio
x60	Net_Income-Exclude_Disposal_Gain_or_Loss/Net_Sales
x61	EPS-Net_Income
x62	Pretax_Income_Per_Share
x63	Retained_Earnings_to_Total_Assets
x64	Total_Income_to_Total_Expenses
x65	Total_Expenses_to_Assets
x66	Net_Income_to_Total_Assets
x67	Gross_Profit_to_Sales
x68	Net_Income_to_Stockholder's_Equity

x69

	One_if_Net_Income_is_Negative_for_the_Last_Two_Years;_Z
ero_Otherwise	
x70	(Inventory_+Accounts_Receivables)/Equity
x71	Total_Asset_Turnover
x72	Accounts_Receivable_Turnover
x73	Days_Receivable_Outstanding
x74	Inventory_Turnover
x75	Fixed_Asset_Turnover
x76	Equity_Turnover
x77	Current_Assets_to_Sales
x78	Quick_Assets_to_Sales
x79	Working_Capital_to_Sales
x80	Cash_to_Sales
x81	Cash_Flow_to_Sales
x82	No-credit_Interval
x83	Cash_Flow_from_Operating/Current_Liabilities
x84	Cash_Flow_to_Total_Assets
x85	Cash_Flow_to_Liability
x86	CFO_to_Assets
x87	Cash_Flow_to_Equity
x88	Realized_Gross_Profit_Growth_Rate
x89	Operating_Income_Growth
x90	Net_Income_Growth
x91	Continuing_Operating_Income_after_Tax_Growth
x92	Net_Income-Excluding_Disposal_Gain_or_Loss_Growth
x93	Total_Asset_Growth
x94	Total_Equity_Growth
x95	Return_on_Total_Asset_Growth