

Data Mining in Labor Productivity

Omar Danilo Castrillón, Ph. D, Jaime Alberto Giraldo, Ph. D, and Jaime Antero Arango, Ph. D
Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial,
Campus La Nubia Bloque Q piso 2, Manizales - Colombia. Odcastrillong@unal.edu.co, jaiagiraldog@unal.edu.co,
jaarangom@unal.edu.co

Abstract— The fundamental objective of this article is to determine, through data mining analysis, the most influential variables in the labor productivity of a person. Parting from a previously designed database and with use of the Weka platform, it is determined through a statistical selection process, with an effectiveness greater than 82.15%, the most important variables. They are: Incentives, standard minute value, production target and number of workers.

Keywords—Data mining, labor productivity, weka, statistical selection, Database

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LACCEI2022.1.1.19>

ISBN: 978-628-95207-0-5 **ISSN:** 2414-6390

Minería de Datos en la Productividad Laboral

Omar Danilo Castrillón, Ph. D, Jaime Alberto Giraldo, Ph. D, and Jaime Antero Arango, Ph. D
Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Industrial,
Campus La Nubia Bloque Q piso 2, Manizales - Colombia. Odcstrillong@unal.edu.co, jaigiraldog@unal.edu.co,
jaarangom@unal.edu.co

Resumen— El objetivo fundamental de este artículo es determinar por medio de un análisis de minería de datos, las variables más influyentes en la productividad laboral de una persona. Partiendo de una Base de datos previamente diseñada y Por medio de la plataforma Weka se logra mediante un proceso selección estadística determinar con una efectividad superior al 82.15% las variables más importantes. Las cuales son: Incentivos, valor minuto estándar, objetivo de producción y número de trabajadores.

Keywords—Minería de datos, productividad laboral, weka, selección estadística, Base de datos

Abstract— The fundamental objective of this article is to determine, through data mining analysis, the most influential variables in the labor productivity of a person. Parting from a previously designed database and with use of the Weka platform, it is determined through a statistical selection process, with an effectiveness greater than 82.15%, the most important variables. They are: Incentives, standard minute value, production target and number of workers.

Keywords—Data mining, labor productivity, weka, statistical selection, Database

I. INTRODUCCIÓN

La competitividad de una empresa, esta fuertemente unida a la productividad de sus trabajadores. Este aspecto tiene especial incidencia en algunas empresas como las pertenecientes al sector de las confecciones [1], donde las técnicas basadas en inteligencia artificial permiten lograr resultados bastante buenos. Igualmente, en algunas otras empresas [2] como las aeronáuticas los modelos predictivos, de simulación y aprendizaje automático han permitido optimizar su sistema de producción. En este mismo sentido, estas técnicas pueden ser usadas en campos tan diversos como la predicción de la productividad de petrolero [3]. Así mismo, por medio de modelos estocásticos lineales, algunos autores [4] han logrado predecir eficientemente el comportamiento del estado de productividad laboral en función del capital de trabajo y costo. Igualmente, en este mismo campo de la inteligencia artificial, se encuentran estudios [5] basados en técnicas de Machine Learning las cuales desarrollan metodologías que permiten eliminar las tareas repetitivas y facilitar el análisis de grandes cantidades de datos, lo cual conduce a una mejora en la productividad. Técnicas que cada día son empleadas por un mayor número de empresas. Estudios similares pueden ser encontrados en [6,7,8].

Como se encuentra en la revisión literaria las técnicas de inteligencia artificial han sido empleadas ampliamente en el análisis de la productividad de las empresas con resultados bastantes buenos, en este sentido, en este trabajo se predice la productividad de los trabajadores, basados en técnicas de minería de datos.

En esta investigación se parte de una base de datos previamente creada perteneciente a un repositorio de la Universidad de California, la cual puede ser encontrada en: (<http://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees>). Esta base de datos inicial consta de 14 variables dependientes y una variable independiente, productividad, la cual se encuentra en un rango de 0 a 1.

Inicialmente, como se ilustra en la sección siguiente la base de datos es modificada y por medio de la plataforma de aprendizaje automático y minería de datos denominada Weka [9], se realiza un proceso de predicción estadística de la variable dependiente. Este proceso esta basado en el algoritmo de clasificación J48 de comportamiento similar a un algoritmo bayesiano [10].

Así, el objetivo fundamental de este trabajo es predecir por medio de técnicas de inteligencia artificial, específicamente técnicas de minería de datos, las variables más influyentes en la productividad laboral de un trabajador. Si bien, como se muestra en la bibliografía, existen diversos trabajos que enfocan este tema desde las áreas de la inteligencia artificial, se resalta que existen muy pocos trabajos o nulos, que desarrollan esta predicción de competitividad por medio de una plataforma libre de minería de datos, como lo es la plataforma de aprendizaje automático y minería de datos denominada WEKA. Este aspecto permite que la metodología acá propuesta pueda ser aplicada en un amplio número de empresas y países, redundando en un gran beneficio no solo para la comunidad académica, sino también para toda la población en general.

Finalmente, en este punto se expresa que para su presentación este trabajo ha sido estructurado en las siguientes partes: II) Materiales y Métodos. III) Resultados. IV) Discusiones V) Conclusiones. Finalmente se incluyen los respectivos agradecimientos y Referencias.

II. MATERIALES Y MÉTODOS

El Desarrollo de la metodología parte de una base de datos previamente definida, la cual puede ser encontrada en [1,6]: (<http://archive.ics.uci.edu/ml/datasets/Productivity+Prediction+of+Garment+Employees>). La base de datos que se ilustra en la Tabla I, es una adaptación de la base de datos mostrada en el vínculo descrito en este párrafo.

TABLA I
BASE DE DATOS INICIAL.

Atributo	Descripción
Semana	1,2,3,4,5
Departamento	1,2
Día	Lu, Ma, Mi, Ju, Sa, Do
Equipo	Número de equipos
Objetivo Productividad	
Svm	Valor minuto estándar
Wip	Trabajo en progreso
OT	Sobre tiempo en minutos por equipo
In	Incentivos
Idt	Tiempo muerto
Idm	Trabajadores parados por interrupción en la producción
Cam	1, 2, 3 Numero de cambios en el estilo de producción.
Numt	Numero de trabajadores por equipo
Prod	M, R, B, MB, E (Porcentaje de productividad) -Variable dependiente

La metodología descrita en esta sección consta de los siguientes pasos y se basa en la metodología propuesta en [11]: A) Estructuración de la base de datos. B) Creación del archivo Arff C) Predicción de la variable dependiente D) Selección de atributos E) Nueva predicción variable dependiente F) Interpretación del árbol de clasificación.

A. Estructuración de la Base de Datos

Tomando como referencia la Tabla I, se estructura el archivo de base de datos. Este archivo el cual consta de 1197 registros y 14 atributos, 13 de ellos independientes y uno dependiente.

B. Creación del archivo Arff

En esta sección de la metodología se definen el respectivo encabezado y detalle del archivo Arff, el cual es interpretado por la plataforma de aprendizaje automático y minería de datos denominada WEKA, por medio del algoritmo J48. Este archivo Arff consta de dos partes que se denominan encabezado y detalle.

C. Predicción de la variable dependiente

Definido el archivo Arff (paso anterior), el mismo, es interpretado por medio del algoritmo J48, desde la plataforma Weka. En esta interpretación del algoritmo se usa el mismo set de datos para entrenar y validar el sistema.

D. Selección de atributos

De forma análoga, por medio de la opción denominada selección de atributos en la plataforma weka, se realiza este proceso. La selección de atributos se hace por medio del evaluador Chi Cuadrado y el método denominado *Ranker*.

E. Nueva predicción variable dependiente

Con el conjunto de atributos más influyentes seleccionados en el paso anterior, nuevamente se realiza el proceso de clasificación por medio del algoritmo J48 de la plataforma Weka. De forma igual al paso C de esta sección, en este proceso se emplea el mismo set de datos para entrenar y validar el sistema.

F. Interpretación del árbol de clasificación

Por medio del árbol de clasificación, obtenido en el paso anterior, se interpretan las principales rutas que llevan desde el nodo inicial hasta las hojas del árbol. Las hojas del árbol de clasificación, representan las posibles predicciones de las variables dependientes. Igualmente, la secuencia de nodos previos a cada hoja, encontrados en cada ruta, contiene los estados de las variables independientes que influyen en el comportamiento de la variable dependiente, es decir en su predicción.

III. RESULTADOS

Como consecuencia de desarrollar la metodología propuesta en cada uno de los pasos descritos en la sección anterior, se obtienen los siguientes resultados:

A. Estructuración de la Base de Datos

La Tabla II ilustra la base de datos definida, con base en la Tabla I. Por razones de espacio solo se muestra una pequeña fracción de esta base de datos. En la Tabla II, cada uno de los campos de izquierda a derecha, son correspondientes con cada uno de los nombres los campos, de abajo hacia arriba definidos en la Tabla I.

TABLA II
FRAGMENTO DE LA BASE DE DATOS.

1	3	2	Do	5.0	0.75	4.15	0.0	240.0	0.0	0.0	0.0	0	2.0	B
2	1	2	Sa	6.0	0.5	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
3	1	2	Lu	5.0	0.6	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
4	2	2	Do	6.0	0.75	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
6	1	2	Ju	8.0	0.75	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B

B. Creación del archivo Arff

Definida la base de datos y su estructura, se define el respectivo archivo Arff. Este archivo consta de dos partes cabeza y detalle, las cuales con ilustradas en las tablas III y IV respectivamente.

TABLA III
ENCABEZADO ARCHIVO ARFF.

Atributo	Tipo
@RELATION	produc_weka
@ATTRIBUTE Semana	{1,2,3,4,5}
@ATTRIBUTE Dep	{1,2}
@ATTRIBUTE Dia	{Lu, Ma, Mi, Ju Sa, Do}
@ATTRIBUTE Equipo	Integer
@ATTRIBUTE Obj_produ	Real
@ATTRIBUTE Smv	Real
@ATTRIBUTE Wip	Real
@ATTRIBUTE OT	Integer
@ATTRIBUTE In	Integer
@ATTRIBUTE Idt	Real
@ATTRIBUTE Idm	Integer
@ATTRIBUTE cam	{0,1,2}
@ATTRIBUTE numt	integer
@ATTRIBUTE prod	{M, R, B, MB, E}

TABLA IV
DETALLE ARCHIVO ARFF.

1	3	2	Do	5.0	0.75	4.15	0.0	240.0	0.0	0.0	0.0	0	2.0	B
2	1	2	Sa	6.0	0.5	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
3	1	2	Lu	5.0	0.6	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
4	2	2	Do	6.0	0.75	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B
6	1	2	Ju	8.0	0.75	2.9	0.0	960.0	0.0	0.0	0.0	0	8.0	B

C. Predicción de la variable dependiente

Por medio del algoritmo J48 se logra un nivel de éxito superior al 84%. El resumen de la clasificación obtenida en esta instancia se ilustra en la Tabla V. Igualmente, las tablas VI y VII ilustran las matrices de precisión y confusión respectivamente.

TABLA V
RESUMEN CLASIFICACIÓN

Nombre	Valor	% Exito
Correctly Classified Instances	1009	84.2941 %
Incorrectly Classified Instances	188	15.7059 %
Kappa statistic	0.7584	
Mean absolute error	0.0935	
Root mean squared error	0.2162	
Relative absolute error	35.0641 %	
Root relative squared error	59.2396 %	
Total Number of Instances	1197	

TABLA VI
MATRIZ DE PRECISIÓN

TP Rate	FP Rate	Preci	Recall	F-Meas	ROC	Class
0.58	0.013	0.77	0.58	0.662	0.97	M
0.667	0.026	0.763	0.667	0.711	0.945	R
0.819	0.063	0.874	0.819	0.846	0.951	B
0.955	0.144	0.84	0.955	0.894	0.961	MB
0.73	0.001	0.964	0.73	0.831	0.984	E
0.843	0.089	0.842	0.843	0.839	0.957	Prom

TABLA VII
MATRIZ DE PRECISIÓN

a	b	c	d	e	classified as
47	15	7	11	1	a = M
9	90	22	14	0	b = R

a	b	c	d	e	classified as
3	10	340	62	0	c = B
1	3	20	505	0	d = MB
1	0	0	9	27	e = E

D. Selección de atributos

Como resultado de la ejecución de este paso en la metodología, se encuentra la siguiente clasificación en importancia de los atributos. Esta clasificación es ilustrada en la Tabla VIII.

TABLA VIII
CLASIFICACIÓN ATRIBUTOS

Atributo	Valor	%	% Acumulado
IN	1198,0118	0,393044635	0,393044635
SMV	550,9139	0,180744257	0,573788892
Obj_produ	457,5167	0,150102432	0,723891324
Numt	297,7834	0,097697007	0,821588332
Equipo	132,2518	0,043389273	0,864977604
Cam	87,4976	0,02870628	0,893683884
WIP	67,4647	0,022133871	0,915817755
IDM	60,1526	0,019734911	0,935552666
IDT	60,1526	0,019734911	0,955287578
OT	58,7504	0,019274877	0,974562454
Semana	37,517	0,012308606	0,986871061
Dep	21,3105	0,006991565	0,993862626
Dia	18,7069	0,006137374	1

E. Nueva predicción variable dependiente

En La Tabla VIII, se identifica que los cuatro primeros atributos contienen el 82.15% de la información, en consecuencia, estos atributos son seleccionados, con el fin de realizar un nuevo proceso de clasificación y predicción de la variable dependiente, solo con estos cuatro atributos. La Tabla IX, ilustra el resumen de esta clasificación, donde se logra un porcentaje de éxito superior al 77.6%. La Tabla X y XI ilustran las respectivas matrices de precisión y confusión

TABLA IX
CLASIFICACIÓN CON ATRIBUTOS SELECCIONADOS

Nombre	Valor	%Exito
Correctly Classified Instances	929	77.6107 %
Incorrectly Classified Instances	268	22.3893 %
Kappa statistic	0.6506	
Mean absolute error	0.1304	
Root mean squared error	0.2554	
Relative absolute error	48.8937 %	
Root relative squared error	69.9531 %	
Total Number of Instances	1197	

TABLA X
MATRIZ DE PRECISIÓN CON ATRIBUTOS SELECCIONADOS

TP Rate	FP Rate	Preci	Recall	F-Meas	ROC	Class
0.407	0.015	0.66	0.407	0.504	0.905	M
0.511	0.039	0.627	0.511	0.563	0.882	R
0.73	0.066	0.854	0.73	0.787	0.907	B
0.945	0.235	0.761	0.945	0.843	0.927	MB
0.649	0.001	0.96	0.649	0.774	0.955	E
0.776	0.132	0.777	0.776	0.767	0.915	prom

TABLA XI
MATRIZ DE CONFUSIÓN CON ATRIBUTOS SELECCIONADOS

a	b	c	d	e	classified as
33	16	13	19	0	a = M
12	69	20	34	0	b = R
3	18	303	91	0	c = B
2	7	19	500	1	d = MB
0	0	0	13	24	e = E

F. Interpretación del árbol de clasificación

Por medio de un proceso similar al realizado en el paso C de la metodología, se vuelve a generar el respectivo árbol de clasificación, con los 4 atributos seleccionados en el paso inmediatamente anterior: IN, SVM, Obj_produ, Numt. La Figura I y la Figura II ilustran este árbol de clasificación. En este árbol se muestran la relación de variables dependientes (nodos) las cuales conducen a un estado específico de la variable dependiente, (hojas).

```

In <= 49
| numt <= 31.5
| | Smv <= 3.9
| | | numt <= 8
| | | | Obj_produ <= 0.75: R (49.0/27.0)
| | | | Obj_produ > 0.75
| | | | | numt <= 6: M (2.0/1.0)
| | | | | numt > 6: B (47.0/13.0)
| | | | numt > 8: MB (25.0/9.0)
| | | Smv > 3.9
| | | | Smv <= 4.3: MB (303.0/121.0)
| | | | Smv > 4.3
| | | | | numt <= 8
| | | | | Obj_produ <= 0.75: B (21.0/10.0)
| | | | | Obj_produ > 0.75: M (5.0/2.0)
| | | | numt > 8: MB (50.0/14.0)
| numt > 31.5
| | Obj_produ <= 0.75
| | | Obj_produ <= 0.5
| | | | Obj_produ <= 0.4
| | | | | In <= 26: M (16.0/3.0)
| | | | | In > 26: R (3.0/1.0)
| | | | Obj_produ > 0.4: R (32.0/8.0)
| | | Obj_produ > 0.5
| | | | In <= 21
| | | | | Smv <= 20.2
| | | | | | Obj_produ <= 0.7
| | | | | | | Smv <= 16.1: B (2.0)
| | | | | | | Smv > 16.1: M (6.0/1.0)
| | | | | | Obj_produ > 0.7: B (8.0/1.0)
| | | | | Smv > 20.2
| | | | | | numt <= 57.5
| | | | | | | Smv <= 21.25: R (4.0)
| | | | | | | Smv > 21.25
| | | | | | | | Smv <= 21.98: B (2.0)
| | | | | | | | Smv > 21.98
| | | | | | | | Smv <= 26.66: M (14.0/7.0)

```

```

| | | | | | | | Smv > 26.66
| | | | | | | | | numt <= 56.5: R (2.0/1.0)
| | | | | | | | | numt > 56.5
| | | | | | | | | | Obj_produ <= 0.65: M (3.0/1.0)
| | | | | | | | | | Obj_produ > 0.65: B (4.0/1.0)
| | | | | | | | numt > 57.5
| | | | | | | | | Obj_produ <= 0.65: R (6.0/2.0)
| | | | | | | | | Obj_produ > 0.65: B (23.0/7.0)
| | | | | | | | | In > 21: B (176.0/7.0)
| | | | | | | | Obj_produ > 0.75
| | | | | | | | | In <= 35
| | | | | | | | | | numt <= 58.5
| | | | | | | | | | | numt <= 55.5: B (7.0/3.0)
| | | | | | | | | | | numt > 55.5: R (11.0/1.0)
| | | | | | | | | | | numt > 58.5
| | | | | | | | | | | numt <= 59: MB (6.0/3.0)
| | | | | | | | | | | numt > 59: M (2.0/1.0)
| | | | | | | | | In > 35
| | | | | | | | | | Smv <= 29.4: MB (18.0)
| | | | | | | | | | Smv > 29.4
| | | | | | | | | | | In <= 38: MB (3.0/1.0)
| | | | | | | | | | | In > 38: R (3.0/1.0)
In > 49
| | | | | | | | In <= 88
| | | | | | | | | Obj_produ <= 0.75
| | | | | | | | | | In <= 63
| | | | | | | | | | | In <= 50: B (21.0)
| | | | | | | | | | | In > 50
| | | | | | | | | | | | In <= 55
| | | | | | | | | | | | | Smv <= 24.26: MB (5.0)
| | | | | | | | | | | | | Smv > 24.26: B (3.0/1.0)
| | | | | | | | | | | | | In > 55: B (31.0/7.0)
| | | | | | | | | | | In > 63: MB (26.0/1.0)
| | | | | | | | | | Obj_produ > 0.75
| | | | | | | | | | | Smv <= 28.08: MB (180.0/4.0)
| | | | | | | | | | | Smv > 28.08
| | | | | | | | | | | | Smv <= 29.12: B (5.0/1.0)
| | | | | | | | | | | | Smv > 29.12: MB (23.0/2.0)
In > 88
| | | | | | | | | | Smv <= 12.52
| | | | | | | | | | | Smv <= 2.9: MB (3.0/1.0)
| | | | | | | | | | | Smv > 2.9
| | | | | | | | | | | | Obj_produ <= 0.65: M (2.0/1.0)
| | | | | | | | | | | | Obj_produ > 0.65: B (5.0/1.0)
| | | | | | | | | | Smv > 12.52
| | | | | | | | | | | In <= 100
| | | | | | | | | | | | In <= 90: E (3.0/1.0)
| | | | | | | | | | | | In > 90: MB (13.0/1.0)
| | | | | | | | | | | In > 100
| | | | | | | | | | | | numt <= 56: MB (2.0)
| | | | | | | | | | | | numt > 56: E (22.0)

```

FIGURA II
ÁRBOL DE CLASIFICACIÓN LINEAL

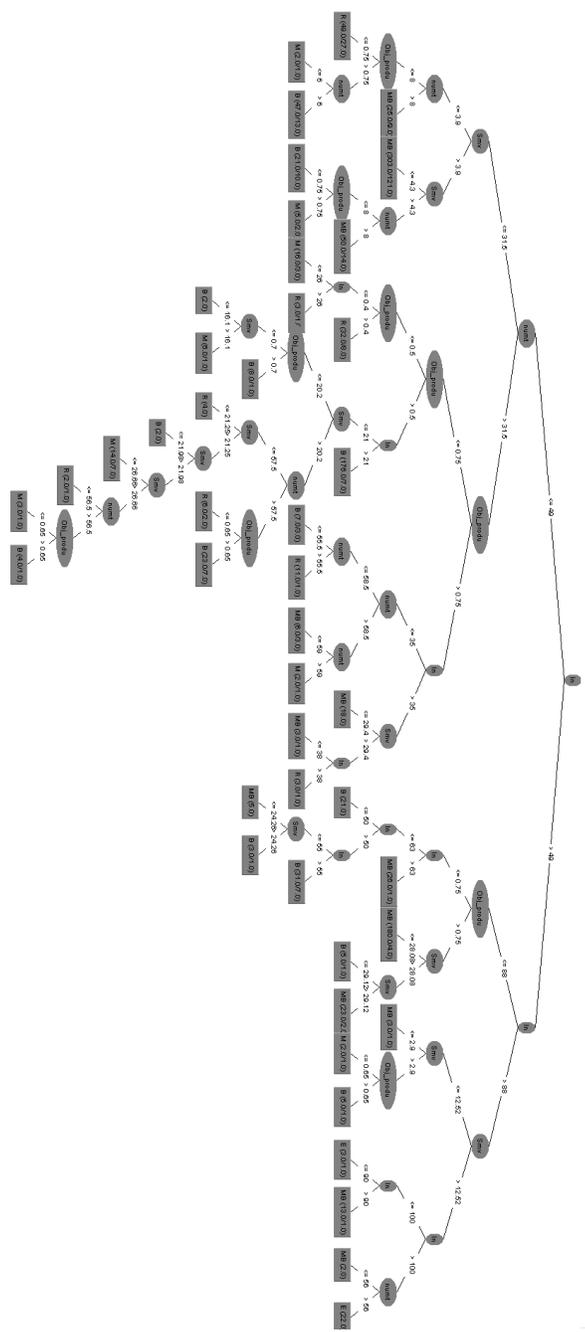


FIGURA I
RELACIÓN VARIABLES DEPENDIENTE E INDEPENDIENTE

IV. DISCUSIONES

Un estudio de la literatura relacionada, muestra que no existe un análisis el cual Identifique por medio un programa de minería de datos libre, como el Weka las principales variables influyente en la productividad de una persona. En este caso se

identifican 4 variables predominantes: Incentivos (IN), valor por minuto (SVM), objetivos de producción (Obj_prod) y número de trabajadores. En este mismo sentido, el árbol de la Figura 1, muestra que cuando los incentivos son muy buenos, la productividad siempre es Buena, Muy buena o Excelente. No obstante, si los incentivos no son los adecuados la productividad empieza a ser mala o regular solo en los siguientes casos: a) cuando el número de trabajadores es bajo, el valor por minuto es bajo y los objetivos de productividad son bajos. b) aunque el numero de trabadores se alto, si los objetivos de productividad son bajos la productividad es baja. c) en algunas ocasiones, aunque los incentivos sean bajos, el número de trabajadores sea bajo y los objetivos de producción bajos, la productividad puede ser buena si el valor por minuto es bueno.

V. CONCLUSIONES

Como resultado de todo el proceso se puede concluir que la productividad de una persona está altamente relacionada con los incentivos y remuneración que la persona pueda recibir. No obstante, estos factores, aunque son determinantes no son los únicos influyentes. Como se deduce de los resultados una productividad mala o regular en un trabajador en esencia esta unida a una serie de valores muy bajos de las cuatro variables identificadas en este proceso: Incentivos (IN), valor por minuto (SVM), objetivos de producción (Obj_prod) y número de trabajadores (numt). Siendo estos aspectos los más predominantes que deberán ser abordados las empresas si desean mejorar la productividad de sus trabajadores.

AGRADECIMIENTOS

Se agradece la colaboración a la Universidad Nacional de Colombia Sede Manizales. Igualmente se agradece al repositorio de bases de datos: Dua, D. y Graff, C., UCI Machine Learning Repository encontrado en el siguiente vinculo de internet: [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science, (2019).

REFERENCIAS

[1] Abdullah Al Imran; Md Shamsur Rahim; Tanvir Ahmed, “Mining the productivity data of the garment industry”, *International Journal of Business Intelligence and Data Mining (IJBIDM)*, Vol. 19, No. 3, diciembre 2021.

[2] Tengfei Long, Yuan Li, Jun Chen, “Productivity prediction in aircraft final assembly lines: Comparisons and insights in different productivity ranges”, *Journal of Manufacturing Systems*, Volume 62, pp. 377-389, Octubre 2021.

[3] Yang Wang, Shiqing Cheng, Haiyang Yu, “Big data technique in the reservoir parameters’ prediction and productivity evaluation: A field case in western South China sea”, *Gondwana Research*, Vol. 96, 22 – 36, Abril 2021.

- [4] Amirza Abdenov, Gaukha'r Abdenova, Ayagoz Mukhanova, "Labour Productivity Prediction Estimate in Manufacturing Environment using Regression and Dynamic Models", *Materials Today: Proceedings*, Vol 16, Part 2, pp 254-261, Agosto 2019.
- [5] K. K. Ramachandran, A. Apsara Saleth Mary, J. R. Pitroda, "Machine learning and role of artificial intelligence in optimizing work performance and employee behavior", *Materials Today: Proceedings*, Article in press, Diciembre 2021.
- [6] Imran, A. A., Amin, M. N., Islam Rifat, M. R., y Mehreen, S. "Deep Neural Network Approach for Predicting the Productivity of Garment Employees", 6th International Conference on Control, Decision and Information Technologies, 2019.
- [7] Joseph Nwogu Nkemakolam, Chisaa Onyekachi Igbolekwu, Esther Chinyere Nwogu, Nwabugwu Obioha Nnanna, Abiola John Asaleye, Rasak Bamidele, Ogadimma Arisukwu, "Psychosocial predictors of emotional labour among employees of merged and acquired banks in Ibadan, South-West, Nigeria", *Heliyon*, Vol 7, Nro e07958, 2021.
- [8] Arpit Mallick, Subhra Dhara, Sushant Rath, "Application of machine learning algorithms for prediction of sinter machine productivity", *Machine Learning with Applications*, Vol 6, Nro 100186, 2021.
- [9] Witten, I., Frank, E., y otros dos autores, "Data Mining Practical Machine Learning Tools and Techniques", Morgan and Kaufman publication (Elsevier), ISBN-13: 978-0128042915, Cambridge, USA (2017)
- [10] Valencia, M., Correa, J., y Díaz, F., "Métodos Estadísticos Clásicos y Bayesianos para el Pronóstico de Demanda. Un Análisis Comparativo", *Revista Facultad de Ciencias Universidad Nacional de Colombia*, Vol. 4 Nro. 1, pp. 52 -67, 2015.
- [11] Castrillón, O., "Predicción del Divorcio por Técnicas de Minería de Datos", *Información Tecnológica*, Vol. 32 Nro. 5, pp. 111-120, Octubre 2021.