

# Análisis exploratorio y aplicación de algoritmos de ML en el sector retail usando conjunto de datos de transacciones en línea

Seminario  
Entregable 2

Especialización en Analítica y Ciencia de Datos  
Departamento de Ingeniería de Sistemas  
Universidad de Antioquia, Colombia

Mario E. Otero, Lina M. Beltrán

Repositorio: [GitHub](#)

**Resumen** - En el sector retail, el análisis de datos de transacciones en línea se ha convertido en una fuente crucial de información para las empresas en su búsqueda por comprender el comportamiento de los clientes y mejorar su toma de decisiones. Este ejercicio se centra en el análisis exploratorio y la aplicación de algoritmos de Machine Learning (ML) sobre un conjunto de datos de transacciones en línea. En primer lugar, se lleva a cabo un análisis exploratorio de los datos, que incluye la limpieza y preparación de los conjuntos de datos, a continuación, se aplican diferentes algoritmos de ML de tipo regresión y predictivos. Se emplean algoritmos de aprendizaje supervisado, como las máquinas de vectores de soporte (SVM) y los árboles de decisión, para predecir el monto total por transacción y así buscar identificar patrones de comportamiento y compra en el sector objeto de estudio.<sup>1</sup>

**Índice de Términos** - Algoritmos, Análisis de datos, Análisis de regresión, Aprendizaje automático, Ciencia de datos, Modelos controlados por datos, Retail.

## I. INTRODUCCIÓN

El presente ejercicio se enfoca en la aplicación de distintas técnicas de regresión, incluyendo regresión lineal simple, regresión lineal múltiple y el uso de técnicas de regularización como Ridge y Lasso. La transformación de datos también será considerada como parte integral del proceso de análisis, con el objetivo de mejorar la calidad de las predicciones.

En primer lugar, se explorará la regresión lineal simple, luego la múltiple, seguido de la regresión lineal múltiple con regularización Ridge y Lasso, además de clasificación y regresión con técnicas ML, tales como: Árboles, Random

Forest y SVM, se evalúa además el rendimiento y precisión de Los modelos con una serie de métricas de desempeño como son: RMSE, R2, MSE y MAE.

## II. DESARROLLO DE CONTENIDO

### A. Etapa de limpieza

Inicialmente, se carga e identifica el estado natural del conjunto de datos para comprender rápidamente la estructura y la calidad de los mismos, e iniciar con la definición de los pasos a seguir para la estrategia de limpieza y transformación, a continuación una muestra de la información de este.

```
1 df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1067371 entries, 0 to 1067370
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   Invoice          1067371 non-null object  
1   StockCode       1067371 non-null object  
2   Description     1062989 non-null object  
3   Quantity       1067371 non-null int64   
4   InvoiceDate     1067371 non-null object  
5   Price          1067371 non-null float64  
6   Customer ID    824364 non-null float64  
7   Country        1067371 non-null object  
dtypes: float64(2), int64(1), object(5)
memory usage: 65.1+ MB
```

Fig. 1. Método info() de la biblioteca Pandas de Python que describe información importante sobre el DataFrame en su estado natural.

<sup>1</sup> consultar [entregable 1](#) para ampliar contexto.

Esta etapa incluyó las siguientes acciones:

- Borrado de características irrelevantes.
- Definición de una lista de columnas categóricas y otra de columnas numéricas.
- Limpieza de caracteres no numéricos de la variable “Invoice”.
- Detección y borrado de datos atípicos.
- Generación de nuevas variables.
- Creación de Variables Dummies.

La ejecución de estas acciones derivó en la generación del nuevo conjunto de datos acondicionado para su posterior análisis con técnicas de ML.

```
1 df.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1030626 entries, 0 to 1031376
Data columns (total 21 columns):
#   Column                               Non-Null Count  Dtype
---  -
0   Invoice                               1030626 non-null object
1   Quantity                             1030626 non-null int64
2   Price                                1030626 non-null float64
3   Year                                 1030626 non-null int64
4   Months                              1030626 non-null int64
5   Month_day                           1030626 non-null int64
6   Time_hour                           1030626 non-null int64
7   wk_day                              1030626 non-null int64
8   year_month                          1030626 non-null object
9   TotalSpent                          1030626 non-null float64
10  TotalTransaction                     1030626 non-null float64
11  TotalQuantity                       1030626 non-null int64
12  TotalProductosUnicos                 1030626 non-null int64
13  CodigoUnico                         1030626 non-null int64
14  CountryUnico                        1030626 non-null int64
15  Continent_Asia                      1030626 non-null uint8
16  Continent_Europe                    1030626 non-null uint8
17  Continent_North America             1030626 non-null uint8
18  Continent_North America             1030626 non-null uint8
19  Continent_Oceania                   1030626 non-null uint8
20  Continent_South America             1030626 non-null uint8
dtypes: float64(3), int64(10), object(2), uint8(6)
memory usage: 131.7+ MB
```

Fig. 2. Método info() de la biblioteca Pandas de Python que describe información importante sobre el DataFrame después de la limpieza y transformación.

### B. Etapa de aplicación de algoritmos de ML

En primer lugar, se explorará la **regresión lineal simple**, que es una técnica ampliamente utilizada para establecer relaciones lineales entre una variable dependiente y una variable independiente. Esta técnica permitirá analizar la relación directa entre los montos transaccionados por ticket y una única variable predictora, lo cual proporcionará un punto de partida para el análisis comparativo más exhaustivo.

Para este modelo se definieron las variables de entrada o predictoras y una variable target o de salida, de la siguiente manera:

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1030626 entries, 0 to 1031376
Data columns (total 20 columns):
#   Column                               Non-Null Count  Dtype
---  -
0   Invoice                               1030626 non-null object
1   Quantity                             1030626 non-null int64
2   Price                                1030626 non-null float64
3   Year                                 1030626 non-null int64
4   Months                              1030626 non-null int64
5   Month_day                           1030626 non-null int64
6   Time_hour                           1030626 non-null int64
7   wk_day                              1030626 non-null int64
8   year_month                          1030626 non-null object
9   TotalSpent                          1030626 non-null float64
10  TotalQuantity                       1030626 non-null int64
11  TotalProductosUnicos                 1030626 non-null int64
12  CodigoUnico                         1030626 non-null int64
13  CountryUnico                        1030626 non-null int64
14  Continent_Asia                      1030626 non-null uint8
15  Continent_Europe                    1030626 non-null uint8
16  Continent_North America             1030626 non-null uint8
17  Continent_North America             1030626 non-null uint8
18  Continent_Oceania                   1030626 non-null uint8
19  Continent_South America             1030626 non-null uint8
dtypes: float64(2), int64(10), object(2), uint8(6)
memory usage: 123.8+ MB
```

Fig. 3. Método info() de la biblioteca Pandas de Python que describe información importante sobre el DataFrame con las variables predictoras o de entrada.

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1030626 entries, 0 to 1031376
Data columns (total 1 columns):
#   Column                               Non-Null Count  Dtype
---  -
0   TotalTransaction                     1030626 non-null float64
dtypes: float64(1)
memory usage: 15.7 MB
```

Fig. 4. Método info() de la biblioteca Pandas de Python que describe información fundamental sobre el DataFrame con la variable target o de salida.

Adicionalmente, se crearon los conjuntos de datos de entrenamiento y prueba en una proporción de 70/30 respectivamente. Finalmente, luego de entrenado el modelo se obtuvo las siguientes métricas:

	Variable	RMSETrain	RMSETest	R2Train	R2Test	MSETrain	MSETest	MAETrain	MAETest
0	Invoice	-2107.370208	-2098.425127	0.005834	0.005841	-4.441607e+06	-4.404353e+06	-1189.517071	-1185.862949
1	Quantity	-2113.026162	-2104.051523	0.000490	0.000504	-4.465480e+06	-4.428005e+06	-1191.338150	-1187.725597
2	Price	-2112.358472	-2102.549003	0.001122	0.001919	-4.462660e+06	-4.421538e+06	-1191.069686	-1187.046354
3	Year	-2112.259421	-2103.243128	0.001215	0.001273	-4.462237e+06	-4.424608e+06	-1194.372826	-1190.701185
4	Months	-2070.677179	-2061.069696	0.040143	0.040909	-4.288242e+06	-4.248893e+06	-1205.027910	-1201.311327
5	Month_day	-2105.771293	-2096.783902	0.007338	0.007397	-4.434854e+06	-4.397469e+06	-1191.329686	-1187.704802
6	Time_hour	-2105.688270	-2096.676705	0.007422	0.007501	-4.434527e+06	-4.397033e+06	-1177.744641	-1174.014785
7	wk_day	-2095.407611	-2087.059389	0.017087	0.016570	-4.391317e+06	-4.356721e+06	-1168.338075	-1165.225222
8	year_month	-2113.008687	-2104.020999	0.000506	0.000534	-4.465405e+06	-4.427882e+06	-1193.609243	-1189.953471
9	TotalSpent	-2110.514542	-2100.320873	0.002864	0.004029	-4.454869e+06	-4.412255e+06	-1190.059944	-1185.924266
10	TotalQuantity	-1037.941576	-1038.518483	0.758781	0.756436	-1.077405e+06	-1.078717e+06	-560.753191	-560.502114
11	TotalProductosUnicos	-1249.370700	-1245.590431	0.850580	0.649728	-1.561334e+06	-1.552188e+06	-584.755675	-582.059449
12	CodigoUnico	-2113.437573	-2104.508628	0.000101	0.000071	-4.467220e+06	-4.429936e+06	-1192.438213	-1188.750328
13	CountryUnico	-2113.497594	-2104.502289	0.000044	0.000078	-4.467474e+06	-4.429914e+06	-1191.766851	-1187.980325
14	Continent_Asia	-2113.509084	-2104.532667	0.000033	0.000048	-4.467522e+06	-4.430039e+06	-1191.900045	-1188.133716
15	Continent_Europe	-2113.195374	-2104.184908	0.000330	0.000378	-4.466194e+06	-4.428572e+06	-1191.733701	-1188.000680
16	Continent_North America	-2113.552622	-2104.598559	-0.000008	-0.000015	-4.467706e+06	-4.430313e+06	-1192.340838	-1188.664147
17	Continent_North America	-2113.566140	-2104.622107	-0.000021	-0.000037	-4.467763e+06	-4.430412e+06	-1192.249951	-1188.588730
18	Continent_Oceania	-2112.571955	-2103.620179	0.000918	0.000914	-4.463554e+06	-4.426192e+06	-1191.873106	-1188.214142

Fig. 5. Tabla de resultados para las métricas aplicadas al modelo de regresión lineal simple.

Encontrando que para las variables “TotalQuantity” y “TotalProductosUnicos” la métrica R2 presentó los resultados más altos, un valor alto de R2 indica que el modelo es capaz de explicar una gran parte de la variabilidad de los datos, lo que sugiere que se ajusta bien a los datos. Sin embargo, es valioso tener en cuenta que un alto valor de R2 no garantiza que el modelo sea válido o que las predicciones sean precisas en todos los casos.

El siguiente experimento aplicado al conjunto de datos fue el algoritmo de **regresión Lineal Múltiple**, que extiende el enfoque de la regresión lineal simple al considerar múltiples variables predictoras. Esto permitirá examinar cómo la inclusión de variables adicionales puede mejorar la precisión de las predicciones y brindar una visión más completa de los factores que influyen en los montos transaccionados por ticket en el sector retail.

Para este caso se realiza el escalamiento de los datos con el método MinMaxScaler de la librería sklearn, Este escalador transforma los datos para que se encuentren en el rango específico indicado, en este caso, de -1 a 1. Adicionalmente, se aplica validación cruzada con 10 pliegues utilizando la función cross\_validate de scikit-learn, Esto implica dividir el conjunto de entrenamiento en múltiples partes (10 en este caso) y ejecutar el entrenamiento y la evaluación del modelo en diferentes combinaciones de conjuntos de entrenamiento y validación. Obteniendo los siguientes resultados.

```
-----INICIO PROCESAMIENTO-----
Resultados RMSE
RMSE_TRAIN: -818.619
RMSE_TEST: -850.561
-----
Resultados R2
R2_TRAIN: 0.850
R2_TEST: 0.836
-----
Resultados MSE
MSE_TRAIN: -670143.773
MSE_TEST: -733188.868
-----
Resultados MAE
MAE_TRAIN: -433.151
MAE_TEST: -433.486
-----
PROCESAMIENTO FINALIZADO EXITOSAMENTE!!!
--- 0.3150031487147013 Minutos ---
```

Fig. 6. Tabla de resultados para las métricas aplicadas al modelo de regresión lineal múltiple.

En general, estos resultados indican que el modelo de regresión lineal múltiple tiene un buen ajuste a los datos, con valores de RMSE, R2, MSE y MAE relativamente bajos.

Finalmente, en la primera iteración (notebook iteracion1ML.ipynb)<sup>2</sup> se aplicó una Regresión Lineal Múltiple

con regularización Ridge y Lasso. Obteniendo los siguientes resultados:

```
-----INICIO PROCESAMIENTO-----
Regresión Lineal Normal:
MSE: 669930.3006431934
R^2: 0.8487875815111954

Regresión Ridge:
MSE: 670876.0610185396
R^2: 0.848574110477672

Regresión Lasso:
MSE: 690713.0200782518
R^2: 0.8440966378928334
-----
PROCESAMIENTO FINALIZADO EXITOSAMENTE!!!
-----
--- 0.8063454031944275 Minutos ---
```

Fig. 7. Tabla de resultados para las métricas aplicadas al modelo de regresión lineal múltiple con regularización Ridge y Lasso.

En general, al comparar estos tres modelos, se observa que la regresión lineal simple y la regresión Ridge tienen resultados similares en términos de MSE y R2, mientras que la regresión Lasso tiene un MSE ligeramente mayor y un R2 ligeramente menor. Sin embargo, la regresión Lasso puede proporcionar la ventaja adicional de seleccionar variables importantes y simplificar el modelo. La elección entre estos modelos dependerá de las características específicas del problema y los objetivos de análisis.

IEEE usará para preparar su documento para su publicación. Escriba el nombre de los autores en la etiqueta del disco. Si usted está usando un Macintosh, por favor guarde su archivo en un disco formateado de PC, si es posible. Usted puede usar Zip o discos de CD-ROM para los archivos grandes, o comprimir archivos usando Winzip o Pkzip. También envíe una hoja de papel con la información completa de contacto para todos los autores. Incluya la dirección de correo geográfico completa, números de teléfono, números de facsímil, y direcciones de correo electrónico. Además, si hay varios autores, elija a un autor como el “autor correspondiente.” Éste es el autor a quien se enviará la corrección del documento. Sólo se envían las evaluaciones al autor correspondiente.

### C. Figuras.

Se procesarán todas las tablas y figuras como imágenes. La IEEE no puede extraer las tablas y figuras incluidas en su documento. (Las figuras y tablas que usted inserta en su documento están solo para ayudarle a medir el tamaño de su documento, por conveniencia de los árbitros, y para hacerle fácil a usted distribuir las preimpresiones.) Por tanto, envíe en hojas de papel separadas, versiones agrandadas de las tablas y figuras que aparecen en su documento. Éstas son las imágenes que la IEEE examinará y publicará con su documento.

### D. Archivos electrónicos de Imagen (Opcional)

Usted tendrá mayor control sobre la apariencia de sus figuras

<sup>2</sup>enlace notebook: [iteracion1ML.ipynb](#)

si usted puede preparar los archivos electrónicos de imagen. Si usted no tiene las habilidades de computación requeridas, sólo envíe las impresiones de papel como se describió anteriormente y salte esta sección.

1) *la Manera más fácil*: Si usted tiene un escáner, la mejor manera y más rápida de preparar los archivos de la figura sin color es imprimir sus tablas y figuras en el papel exactamente como usted quiere que ellas aparezcan, explórelas (con el scanner), y luego guárdelas en un archivo en PostScript (PS) o PostScript encapsulado (EPS). Use un archivo separado para cada imagen. Los nombres de los archivos deben tener el formato "fig1.ps" o "fig2.eps."

2) *La manera un poco más difícil*: Usando un scanner como se describe en el numeral 1, guarde las imágenes en formato TIFF. Alto-contraste línea de línea de figuras y tablas deben prepararse con resolución de 600 dpi y salvadas sin compresión. 1 bit por pixel (monocromo), los nombres de los archivos de la forma "fig3.tif" o "table1.tif.". Para obtener un 3,45-en la figura (una-columna de ancho) a 600 dpi la figura requiere un tamaño horizontal de 2070 píxeles. El tamaño típico de los ficheros será del orden de 0,5 MB.

Fotografías y figuras en escala de grises deben prepararse con 220 dpi de resolución y salvadas sin compresión, 8 bits por pixel (escala de grises). Para obtener un 3,45-en la figura (una-columna de ancho) a 220 dpi, la figura debería tener un tamaño horizontal de 759 píxeles.

El color de las figuras se deben preparar con 400 dpi de resolución y guardado sin compresión, 8 bits por pixel (paleta ó 256 color). Para obtener un 3,45-en la figura (una-columna de ancho) a 400 dpi, la figura debería tener un tamaño horizontal de 1380 píxeles.

Para obtener más información sobre los archivos TIFF, por favor vaya a <http://www.ieee.org/organizations/pubs/transactions/informati> on.htm y haga clic en el enlace "Guidelines for Author Supplied Electronic Text and Graphics."

3) Algo mas difícil: Si no dispone de un escáner, puede crear noncolor PostScript figuras por "impresión" a los archivos. En primer lugar, descargar un controlador de impresora PostScript de <http://www.adobe.com/support/downloads/pdrvwin.htm> (para Windows) o desde <http://www.adobe.com/support/downloads/pdrvmac.htm> (por Macintosh) e instalar el "Generic PostScript para imprimir" definición. En Word, pegue la imagen en un nuevo documento. Envía a Imprimir el archivo utilizando el controlador de impresora PostScript. El Nombre de los archivos debe tener el formato "fig5.ps." Utilize fuentes Adobe Tipo 1 al crear su figuras, de ser posible.

4) Otra forma: Usuarios avanzados de la computadora puede convertir figuras y tablas de su formato original al formato TIFF. Algunas convertidores de imágenes útiles son Adobe Photoshop, Corel Draw, y Microsoft Photo Editor, una

aplicación que forma parte de Microsoft Office 97 y Office 2000 (busque C: \ Program Files \ Common Files \ Microsoft Shared \ PhotoEd \ PHOTOED.EXE. (Usted tiene la opción de instalar el photo editor de su disco original de Office)

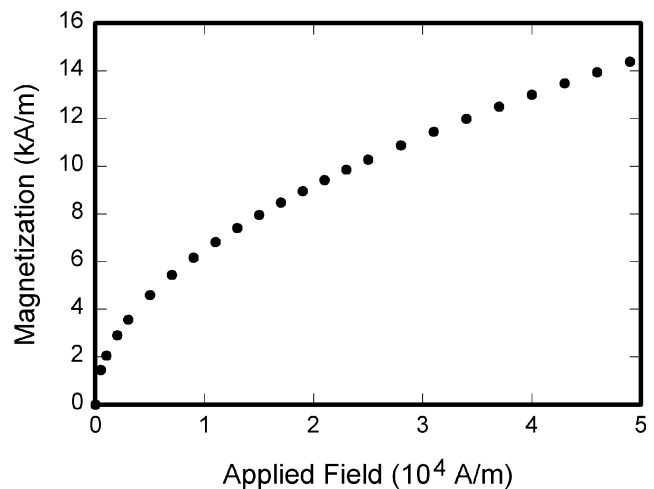


Fig. 1. función de Magnetización. Tenga en cuenta que "Fig." Está abreviado. Hay un espacio después del número de figura, seguido por dos espacios. Es una buena práctica para explicar la importancia de la figura en el subtítulo.

Esta es una manera de crear archivos de imagen TIFF a partir de tablas. En primer lugar, crear su tabla en word. Use líneas horizontales, pero no líneas verticales. Oculte la cuadrícula(tabla | Ocultar cuadrícula).Ejecute el corrector ortográfico de la tabla para eliminar cualquier palabra subrayada en rojo(indica que subraya los errores ortográficos). Ajuste el aumento (Ver | Zoom), de tal manera que usted pueda ver la tabla completa cuando seleccione Ver | Pantalla completa. Mover el cursor fuera de la vista en pantalla. Presione "impresión de pantalla" de su teclado, esto copia la imagen de pantalla al portapapeles de Windows. Abra Microsoft Photo Editor y haga clic en Editar | Pegar como Nueva Imagen. Recortar imagen de la tabla (haga clic en botón "Seleccionar"; seleccionar la parte que desea, luego de Imagen | recortar). Ajuste de las propiedades de la imagen (Archivo | propiedades) a monocromo (1 bit) y 600 pixels por pulgada. Cambiar el tamaño de la imagen (Imagen | cambio de tamaño) para un ancho de 3,45 pulgadas. Guarde el archivo (Archivo | Guardar como) en formato TIFF sin compresión (haga clic en "mas").

La mayoría de los programas de gráficos le permite guardar los gráficos en formato TIFF, pero que a menudo no tienen ningún control sobre la compresión o el número de bits por pixel. Usted debe abrir estos archivos de imagen en un programa como Microsoft Photo Editor y volver a guardarlos sin usar compresión, ya sea 1 ó 8 bits, y ya sea 600 o 220 dpi de resolución (Archivo | propiedades; Imagen | cambio de tamaño). Véase la Sección II-D2 para una explicación de número de bits y resolución. Si su programa de gráficos no puede exportar a TIFF, puede utilizar la misma técnica descrita para las tablas en el párrafo anterior.

No vertical lines in table. Statements that serve as captions for the entire table do not need footnote letters.

\*Gaussian units are the same as cgs emu for magnetostatics; Mx = maxwell, G = gauss, Oe = oersted; Wb = weber, V = volt, s = second, T = tesla, m = meter, A = ampere, J = joule, kg = kilogram, H = henry.

La  
unidad  
del SI  
para la  
fuerza  
del  
campo

Una manera de convertir una figura de Windows Metafile (WMF) a TIFF es pegarlo en Microsoft PowerPoint, guardarla en formato JPG, abrirlo con Microsoft Photo Editor o similar convertidor, y volver a guardarlo como TIFF.

Microsoft Excel le permite guardar la hoja de cálculo en formato "Graphics Interchange Format" (GIF). Para obtener una buena resolución, haga gráficos en Excel muy grandes. A continuación, utilice el "Guardar como HTML" (véase <http://support.microsoft.com/support/kb/articles/q158/0/79.asp>). A continuación, puede convertir a TIFF de GIF con Microsoft Photo Editor, por ejemplo.

No importa como hayan sido convertidas las imágenes, es una buena idea imprimir los archivos TIFF para asegurarse de que nada se perdió en la conversión.

Si modifica este documento para su uso con otras conferencias o revistas IEEE, debe guardarlo como tipo "Word 97-2000 & 6.0/95 - RTF (\*.doc)" a fin de que pueda ser abierto por cualquier versión de Word.

#### E. Formulario de Copyright

Un formulario de IEEE de derechos de autor debe acompañar su presentación final. Usted puede obtener una versión. Pdf, Html o. Doc en <http://www.ieee.org/copyright>. Los autores son responsables de la obtención de cualquier autorización.

### III. PLAN DE TRABAJO PARA LA MONOGRAFÍA

Si usted está usando *Word*, use el Editor de Ecuaciones de Microsoft o el complemento *MathType* (<http://www.mathtype.com>) para las ecuaciones en su documento (Insertar | Objeto | Crear Nuevo | Editor de Ecuaciones de Microsoft o Ecuación MathType).

### IV. CONCLUSIONES

Use SI (MKS) o CGS como unidades primarias. (Se prefieren las unidades del SI.) Pueden usarse las unidades inglesas como unidades secundarias (en paréntesis). **Esto se aplica a los documentos en el almacenamiento de información.** Por ejemplo, escriba "15 Gb/cm<sup>2</sup> (100 Gb/in<sup>2</sup>)." Una excepción es cuando se usan las unidades inglesas como los identificadores en el comercio, como "3½ en la unidad de disco." Evite combinar SI y unidades de CGS, como la corriente en los amperios y el campo magnético en oersteds. Esto lleva a menudo a confusión porque las ecuaciones no cuadran dimensionalmente. Si usted debe usar unidades mixtas, claramente declare las unidades para cada cantidad en una ecuación.

magnético H es A/m. Sin embargo, si usted desea usar unidades de T, o referirse a densidad de flujo magnético *B* o la fuerza del campo magnético simbolizadas como  $\mu_0 H$ . Use un punto en el centro para separar las unidades compuestas, por ejemplo, "A·m<sup>2</sup>."

### V. INDICACIONES ÚTILES

#### A. Figuras y tablas

Debido a que IEEE dará el último formato de su documento, Las figuras grandes y tablas pueden ocupar el espacio de ambas columnas. Ponga los subtítulos de las figuras debajo de las figuras; ponga los títulos de las tablas sobre las tablas. Si su figura tiene dos partes, incluya las etiquetas "(a)" y "(b)" como parte de las obras de arte. Por favor verifique que las figuras y tablas que usted menciona en el texto realmente existan. **Por favor no incluya subtítulos como parte de las figuras. No ponga subtítulos en "cuadros de texto" vinculados a las figuras. No ponga bordes externos en sus figuras.** Use la abreviación "Fig." incluso al principio de una frase. No abrevie "Tabla". Las tablas se numeran con números romanos.

**No use color a menos que sea necesario para la interpretación apropiada de sus figuras.** Las etiquetas de los ejes de las figuras son a menudo una fuente de confusión. Use palabras en lugar de símbolos. Como ejemplo, escriba la cantidad "Magnetización," o "Magnetización M," no sólo "M." Ponga las unidades en los paréntesis. No etiquete los ejes sólo con las unidades. Como en la Fig. por ejemplo, 1 escriba "Magnetización (A/m)" o "Magnetización (A·m<sup>-1</sup>)" no sólo "A/m." No etiquete los ejes con una proporción de cantidades y unidades. Por ejemplo, escriba "Temperatura ( K)," no



“Temperatura /K.”

Los multiplicadores pueden ser sobre todo confusos. Escriba “Magnetización (kA/m)” o “Magnetización (10<sup>3</sup> A/m).” No escriba “Magnetización (A/m) × 1000” porque el lector no sabrá si la etiqueta del eje de arriba en la Fig. 1 significa 16000 A/m o 0.016 A/m. Las etiquetas de la figura deben ser legibles, aproximadamente 8 a 12 puntos.

## B. Referencias

Dentro del texto, numere las citas en paréntesis cuadrados [1], siguiendo el orden en el que aparecen relacionadas en la última sección del artículo, llamada REFERENCIAS (En la sección de REFERENCIAS, las referencias deben estar ordenadas en orden lexicográfico por autor). El punto de la frase sigue los paréntesis [2]. Múltiples referencias [2], [3] son numeradas con los paréntesis separados [1]–[3]. Al citar una sección en un libro, por favor dé los números de página pertinentes [2]. En las frases, simplemente refiérase al número de la referencia, como en [3]. No use “Ref. [3]” o “referencia [3]” excepto al principio de una frase: “la Referencia [3] muestra...”

Numere las notas a pie de página separadamente en los exponentes (Insertar | Referencia | Nota a pie de página). Ponga la nota a pie de página real al final (parte inferior) de la columna en que se cita; no ponga las notas a pie de página en la lista de referencias (notas del final). Use letras para las notas a pie de página en la tabla (ver Tabla I).

Por favor note que las referencias al final de este documento están en estilo referido preferido. **Allí están organizadas por orden alfabético del apellido del autor.** Dé todos los nombres de los autores; no use “et al” a menos que haya seis autores o más. Evite el uso de las iniciales de los nombres de los autores. Escriba apellidos y nombres siempre que sea posible. Documentos que no se han publicado deben citarse como “inédito” [4]. Documentos que se han sometido o se han aceptado para la publicación deben citarse como “sometido a publicación” [4]. Por favor dé afiliaciones y direcciones para las comunicaciones personales [6].

Escriba con mayúscula sólo los primeros términos del título del documento, salvo los nombres propios y símbolos del elemento. Si usted esta corto de espacio, puede omitir los títulos del documento. Sin embargo, los títulos del documento son útiles a sus lectores y se recomiendan fuertemente.

## C. Abreviaciones y Siglas

Defina las abreviaciones y siglas la primera vez que sean usadas en el texto, incluso después de que se hayan definido en la teoría. Las abreviaciones como ACM, IEEE, SI, ac, y dc no tienen que ser definidas. Las abreviaciones que llevan puntos incorporados no deben tener espacios: escriba “C.N.R.S.” no “C. N. R. S.” *No use las abreviaciones en el título* a menos que ellas sean inevitables (por ejemplo, “IEEE”

en el título de este artículo).

## D. Ecuaciones

Numere las ecuaciones consecutivamente con los números de la ecuación en paréntesis contra el margen derecho, como en (1). Primero use el editor de ecuaciones para crear la ecuación. Luego seleccione estilo de “Ecuación”. Presione la tecla tab y escriba el número de la ecuación en los paréntesis. Para hacer sus ecuaciones más compactas, usted puede usar (/), la función exp, o exponentes apropiados. Use los paréntesis para evitar las ambigüedades en los denominadores. Puntúe las ecuaciones cuando sean parte de una frase, como en

$$\int_0^{r_2} F(r, \varphi) dr d\varphi = [\sigma r_2 / (2\mu_0)] \cdot \int_0^\infty \exp(-\lambda |z_j - z_i|) \lambda^{-1} J_1(\lambda r_2) J_0(\lambda r_i) d\lambda. \quad (1)$$

Asegúrese de que los símbolos en su ecuación han estado definidos antes de aparecer la ecuación o inmediatamente enseguida. Ponga en cursiva los símbolos (*T* podría referirse a la temperatura, pero *T* es la unidad tesla). Refiérase a “(1),” no a “Eq. (1)” o “la ecuación (1),” excepto al principio de una oración: “la Ecuación (1) es...”

## E. Otras Recomendaciones

Use un espacio después de los puntos finales y de los dos puntos. Una con guión los modificadores complejos: “campo -cero -refrescando la magnetización.” Evite hacer balancear en el aire los participios, como, “Usando (1), el potencial era calculado.” [No está claro quién o que usó (1).] Escriba en cambio, “El potencial era calculado usando (1),” o “Usando (1), se calcula el potencial.”

Use un cero antes de los puntos decimales: “0.25,” no “.25.” Use “cm<sup>3</sup>,” no “cc.” Indique las dimensiones simplificadas como “0.1 cm ∅ 0.2 cm,” no “0.1 ∅ 0.2 cm<sup>2</sup>.” La abreviación para “segundos” es “s,” no “sec.” No mezcle los nombres completos y abreviaciones de unidades: use “Wb/m<sup>2</sup>” o “webers por metro cuadrado,” no “webers/m<sup>2</sup>.” Al expresar un rango de valores, escriba “7 a 9” o “7-9,” no “7~9.”

Una declaración en paréntesis al final de una frase se puntúa fuera del paréntesis del cierre (así está bien). (Una frase en paréntesis se puntúa dentro de los paréntesis.) En inglés americano, los puntos finales y comas van dentro de las comillas, como “este punto.” Otra puntuación va “afuera”! Evite las contracciones; por ejemplo, escriba “do not” en lugar de “don’t.” La coma consecutiva se prefiere: “A, B, y C” en lugar de “A, B y C.”

Evite el uso de la primera persona singular o plural. Pero si debe escoger entre la voz pasiva y la primera persona, puede escribirlo para usar la voz activa (“yo observé que...” o

“Nosotros observamos que...” en lugar de “fue observado que...”). Recuerde verificar la ortografía. Si su idioma nativo no es inglés, por favor consiga que un colega angloparlante nativo corrija su documento.

## VI. ALGUNOS ERRORES COMUNES

La palabra “data (datos)” es plural, no singular. El subíndice para la permeabilidad del vacío  $\mu_0$  es cero, no un escriba en letras minúsculas la letra “o.” El término para la magnetización residual es “remanente”. Use la palabra “micrómetro” en lugar de “microm.” Un gráfico dentro de un gráfico es una “intercalación,” no una “inserción.” La palabra “alternativamente” se prefiere a la palabra “alternadamente” (a menos que usted realmente quiera decir algo que alterne). Use la palabra “considerando que” en lugar de “mientras” (a menos que usted está refiriéndose a los eventos simultáneos). No use la palabra “esencialmente” para significar “aproximadamente” o “eficazmente.” No use la palabra “asunto” como una alusión para “problema.” Cuando las composiciones no son los símbolos químicos especificados, separados por guiones; por ejemplo, “NiMn” indica la aleación  $\text{Ni}_{0.5}\text{Mn}_{0.5}$  compuesto considerando que “Ni-Mn” indica una aleación de alguna composición  $\text{Ni}_x\text{Mn}_{1-x}$ .

Sea consciente de los diferentes significados de los homófonos “afecta” (normalmente un verbo) y “efecto” (normalmente un sustantivo), “complemento” y “cumplimiento,” “continúo” y “discreto,” “principal” (por ejemplo, “el investigador principal”) y “principio” (por ejemplo, “el principio de medida”). No confunda “implicar” e “inferir.”

Los prefijos como “sub,” “micro,” “multi,” y “ultra” no son palabras independientes; ellas deben unirse a las palabras que ellos modifican, normalmente sin un guión. No hay ningún período después “et” en la abreviación latina “*et al.*” (Además se pone en cursiva). La abreviación “i.e.,” significa “es decir,” y la abreviación “e.g.,” significa “por ejemplo” (estas abreviaciones no se ponen cursiva). Un excelente manual de estilos y fuente de información para escritores de la ciencia es [8]. Una guía general de estilos IEEE, *Información para Autores*, está disponible en <http://www.ieee.org/organizations/pubs/transactions/informati on.htm>

## VII. POLÍTICA EDITORIAL

Presentación de un manuscrito no es necesaria para la participación en una conferencia. No envíe una versión de una

nueva presentación de un documento que usted ha enviado o ha publicado en otra parte. No publique datos o resultados “preliminares”. El autor que remite es el único responsable para estar de acuerdo con todos los coautores y cualquier consentimiento requerido de los patrocinadores antes de enviar un documento (paper). La IEEE rechaza radicalmente la paternidad literaria de cortesía. Es obligación de los autores citar el trabajo previo pertinente.

## VIII. PRINCIPIOS DE PUBLICACIÓN

El contenido de TRANSACTIONS y JOURNALS es revisado y archivado por expertos.

Los autores deben considerar los siguientes puntos:

1) Los documentos técnicos enviados para publicación deben adelantar el estado de conocimiento y deben citar el trabajo previo pertinente.

2) La longitud de un documento enviado debe ser correspondiente con la importancia, o apropiado a la complejidad, del trabajo. Por ejemplo, una extensión obvia de trabajo previamente publicado no podría ser apropiada para la publicación o podría tratarse adecuadamente en sólo unas páginas.

3) Los autores deben convencer al Comité Editorial, por medio de su documento, del mérito académico, científico o técnico del documento; las normas de evaluación son más exigentes cuando se reportan resultados extraordinarios o inesperados.

4) Debido a que la repetición se requiere para el progreso científico, los documentos enviados para publicación deben proporcionar información suficiente para permitirles a los lectores tener acceso a las referencias utilizadas, especialmente si estas son URLs, realizar experimentos similares o cálculos y usar los resultados informados. Aunque no todo necesita ser descubierto, un documento debe contener información nueva, usada y totalmente descubierta. Por ejemplo, la composición química de un espécimen necesita que no se informe si el propósito principal de un documento es introducir una nueva técnica de la medida. *Los autores deben esperar ser desafiados por críticos si los resultados no son soportados por los datos adecuados y los detalles críticos.*

5) Documentos que describen el trabajo en curso o muestran un reciente logro técnico, que sean adecuados para su presentación en una conferencia profesional, pueden no ser apropiadas para su publicación en un TRANSACTIONS or JOURNAL

## IX. CONCLUSIÓN

Una sección de conclusión no es necesaria. Sin embargo esta puede repasar los puntos principales del artículo, no repita el resumen como conclusión. Una conclusión se elabora con base en la importancia del trabajo realizado o en las aplicaciones y extensiones sugeridas.

#### APÉNDICE

Los apéndices, si son necesarios, aparecen antes del reconocimiento.

#### RECONOCIMIENTO

Use el título singular aún cuando tenga que hacer muchos reconocimientos. Evite las expresiones como “Uno de nosotros (S.B.A.) gustaría agradecer...” En cambio, escriba “F. A. agradecimientos del autor...” los reconocimientos a un patrocinador y de apoyo financiero se ponen en la nota a pie de página de la primera página sin numerar.

#### REFERENCES

- [1] G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
- [2] W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.
- [3] H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.
- [4] B. Smith, “An approach to graphs of linear forms (Unpublished work style),” unpublished.
- [5] E. H. Miller, “A note on reflector arrays (Periodical style—Accepted for publication),” *IEEE Trans. Antennas Propagat.*, to be published.
- [6] J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.
- [7] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
- [8] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces(Translation Journals style),” *IEEE Transl. J. Magn.Jpn.*, vol. 2, Aug. 1987, pp. 740–741 [*Dig. 9th Annu. Conf. Magnetism Japan*, 1982, p. 301].
- [9] M. Young, *The Technical Writers Handbook*. Mill Valley, CA: University Science, 1989.
- [10] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Periodical style),” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34–39, Jan. 1959.
- [11] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. Neural Networks*, vol. 4, pp. 570–578, July 1993.
- [12] R. W. Lucky, “Automatic equalization for digital communication,” *Bell Syst. Tech. J.*, vol. 44, no. 4, pp. 547–588, Apr. 1965.
- [13] S. P. Bingulac, “On the compatibility of adaptive controllers (Published Conference Proceedings style),” in *Proc. 4th Annu. Allerton Conf. Circuits and Systems Theory*, New York, 1994, pp. 8–16.
- [14] G. R. Faulhaber, “Design of service systems with priority reservation,” in *Conf. Rec. 1995 IEEE Int. Conf. Communications*, pp. 3–8.
- [15] W. D. Doyle, “Magnetization reversal in films with biaxial anisotropy,” in *1987 Proc. INTERMAG Conf.*, pp. 2.2-1–2.2-6.
- [16] G. W. Juette and L. E. Zeffanella, “Radio noise currents in short sections on bundle conductors (Presented Conference Paper style),” presented at the IEEE Summer power Meeting, Dallas, TX, June 22–27, 1990, Paper 90 SM 690-0 PWRs.
- [17] J. G. Kreifeldt, “An analysis of surface-detected EMG as an amplitude-modulated noise,” presented at the 1989 Int. Conf. Medicine and Biological Engineering, Chicago, IL.

- [18] J. Williams, “Narrow-band analyzer (Thesis or Dissertation style),” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993.
- [19] N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.
- [20] J. P. Wilkinson, “Nonlinear resonant circuit devices (Patent style),” U.S. Patent 3 624 12, July 16, 1990.
- [21] *IEEE Criteria for Class IE Electric Systems* (Standards style), IEEE Standard 308, 1969.
- [22] *Letter Symbols for Quantities*, ANSI Standard Y10.5-1968.
- [23] R. E. Haskell and C. T. Case, “Transient signal propagation in lossless isotropic plasmas (Report style),” USAF Cambridge Res. Lab., Cambridge, MA Rep. ARCRL-66-234 (II), 1994, vol. 2.
- [24] E. E. Reber, R. L. Michell, and C. J. Carter, “Oxygen absorption in the Earth’s atmosphere,” Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (420-46)-3, Nov. 1988.
- [25] (Handbook style) *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, 1985, pp. 44–60.
- [26] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.
- [27] (Basic Book/Monograph Online Sources) J. K. Author. (year, month, day). Title (edition) [Type of medium]. Volume(issue). Available: <http://www.URL>
- [28] J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: <http://www.atm.com>
- [29] (Journal Online Sources style) K. Author. (year, month). Title. Journal [Type of medium]. Volume(issue), paging if given. Available: <http://www.URL>
- [30] R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. *IEEE Trans. Plasma Sci.* [Online]. 21(3). pp. 876–880. Available: <http://www.haleyon.com/pub/journals/21ps03-vidmar>

**Biografía Autor(es)** (M’76-SM’81-F’87) y los otros autores pueden incluir las biografías al final de los documentos (papers) regulares. Por favor incluyan nombres y apellidos con los cuales puedan ser identificados al registrar sus artículos (Son registrados en la base de Publindex en Colciencias, entre otras). Si ha enviado documentos antes, no debe suponer que el Comité Editorial conoce o puede decidir cuál es o cuáles son los autores del artículo. *Cada envío debe ser completo en todos sus datos*. El primer párrafo debe contener la filiación institucional (por ejemplo, profesor asociado, Facultad de Ingeniería de Sistemas, Universidad El Bosque). Los grados deben listarse con el tipo de grado, en qué campo, en que institución, ciudad, estado o país.

El segundo párrafo usa el pronombre de la persona (él o ella) y no el apellido o nombre del autor. Lista la experiencia académica y laboral. Se ponen en mayúscula los títulos del trabajo. Pueden listarse cargos anteriores. Información que involucra las publicaciones anteriores puede ser incluida. Intente no listar más de tres libros o artículos publicados. El formato para listar a publicadores de un libro dentro de la biografía es: el título de libro (la ciudad, estado: el nombre del publicador, año) similar a una referencia. Los intereses de investigaciones actuales y anteriores terminan el párrafo.

El tercer párrafo empieza con el título del autor y apellido (por ejemplo, Dr. Smith, Prof. Jones, Sr. Kajor, Ms. Hunter). Finalmente, liste cualquier premio por trabajos y publicaciones. Proporcionar una fotografía es requisito para publicar su artículo: la biografía se dentará alrededor de ella.



La fotografía se pone en la esquina superior izquierda de la biografía. Se quitarán las aficiones personales de la biografía.