

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

Tecnologías de la información



Extracción de Conocimiento en Bases de Datos

V.1. Reporte de investigación de técnicas de visualización

Docente

Enrique Mascote

Alumno

Myriam Raquel Almuina Orozco

IDGS 91N

Sábado, 29 de noviembre del 2025

1. Introducción

La visualización de información es una etapa fundamental dentro del proceso de extracción de conocimiento, ya que permite transformar datos complejos en representaciones gráficas que facilitan su comprensión, análisis y comunicación. A través de gráficos, diagramas y paneles interactivos, los patrones ocultos, tendencias y relaciones entre variables pueden identificarse con claridad.

El presente reporte describe las principales técnicas de visualización de datos, sus características, aplicaciones y ventajas, así como su relación con los procesos analíticos utilizados en minería de datos.

2. Técnicas de visualización y representación de información

A continuación se presentan las técnicas más comunes y utilizadas en ciencia de datos, análisis estadístico y presentación ejecutiva de resultados.

2.1 Gráficos básicos (visualización univariada y bivariada)

Histogramas

Representan la distribución de frecuencias de una variable continua.

Sirven para identificar valores extremos, dispersión y forma de distribución (normal, sesgada, multimodal).

Diagramas de barras

Útiles para comparar categorías o grupos.

Son comunes en reportes ejecutivos y tableros de control.

Gráfico de líneas

Se emplea para representar cambios en el tiempo (series temporales).

2.2 Visualizaciones multivariadas

Mapas de calor (Heatmaps)

Representan valores mediante colores.

Sirven para analizar correlaciones, matrices de distancia y patrones intensivos.

Diagramas de dispersión (Scatter plots)

Permiten visualizar la relación entre dos variables y detectar patrones, clusters o outliers.

Bubble charts

Extienden el scatter plot agregando un tercer atributo mediante el tamaño de la burbuja.

2.3 Técnicas avanzadas para grandes volúmenes de datos

Gráficos de cajas (Boxplots)

Muestran la dispersión, cuartiles y valores atípicos de una variable.

Violin plots

Combinan boxplot + densidad, siendo útiles para comparar distribuciones entre grupos.

Mapas geoespaciales

Representan datos sobre mapas, integrando coordenadas y valores regionales.

Ejemplos: mapas de calor geográfico, coropletas y puntos geolocalizados.

2.4 Visualizaciones interactivas y dashboards

Dashboards

Permiten monitorear indicadores clave en tiempo real mediante gráficos dinámicos.

Herramientas comunes:

- Power BI
- Tableau
- Looker Studio
- Grafana

Plotly y Bokeh (Python)

Bibliotecas que permiten interacción: zoom, hover, filtros, animaciones.

2.5 Visualización en minería de datos y machine learning

Matriz de confusión

Muestra el desempeño de modelos de clasificación.

Curva ROC y AUC

Permite evaluar modelos binarios respecto a sensibilidad y especificidad.

Gráficas de varianza explicada (PCA)

Se usan para observar cuánta información conserva cada componente principal.

T-SNE y UMAP (proyecciones 2D/3D)

Métodos avanzados de reducción de dimensionalidad que permiten visualizar clusters complejos.

3. Relación entre visualización y extracción de conocimiento

La visualización de datos cumple un papel crucial en todas las etapas del proceso KDD (Knowledge Discovery in Databases):

Etapa	Uso de visualización
Selección de datos	Identificar variables relevantes
Limpieza de datos	Detectar valores extremos o inconsistencias
Transformación	Observar distribuciones antes/después de normalizar
Minería de datos	Visualizar resultados de clustering, clasificación o regresión
Interpretación	Tablas, dashboards, gráficos ejecutivos
Comunicación	Presentación de hallazgos mediante infografías y reportes

La visualización permite validar supuestos, detectar problemas en los datos y comunicar descubrimientos a usuarios no técnicos.

4. Conclusiones

Las técnicas de visualización son herramientas esenciales para transformar datos en conocimiento accionable. Su uso adecuado facilita identificar patrones, interpretar resultados de modelos y comunicar información de manera clara y comprensible.

Desde gráficos básicos hasta visualizaciones avanzadas e interactivas, cada técnica aporta una perspectiva distinta que ayuda en el análisis exploratorio y en la presentación final de hallazgos. Una correcta elección de visualización mejora la toma de decisiones y fortalece todo el proceso de extracción de conocimiento.

5. Referencias (APA)

Kelleher, J. D., & Tierney, B. (2018). *Data Science*. MIT Press.

Kirk, A. (2016). *Data Visualisation: A Handbook for Data Driven Design*. Sage.

Few, S. (2009). *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press.

Ware, C. (2013). *Information Visualization: Perception for Design* (3rd ed.). Morgan Kaufmann.

The scikit-learn developers. (2024). *Visualization and model evaluation*.
<https://scikit-learn.org/>