

Técnicas de Minería de Datos

SERIES DE TIEMPO

Equipo 1:

- Jesús Guillermo Félix Barreras 1794656
- Vanessa Stefanía Linares Núñez 1723237
- Dayan Ramirez Vargas 1749744
- Erick Alejandro Rodríguez Estrada 1803744



Prof. Mayra Cristina Berrones Reyes
Grupo: 002

¿Qué son las series de tiempo?

Una serie temporal es una sucesión de observaciones de una variable realizadas a intervalos regulares de tiempo.

El objetivo fundamental de su estudio es conocer el comportamiento de una variable a través del tiempo para, a partir de este, y bajo el supuesto de que no van a producirse cambios estructurales, poder realizar predicciones.



Supuestos

1. Se considera que existe una cierta **estabilidad en la estructura** del fenómeno estudiado.
2. Los datos deben ser **homogéneos en el tiempo**, o lo que es lo mismo, se debe mantener la definición y la medición de la magnitud objeto de estudio.



Técnicas de Predicción

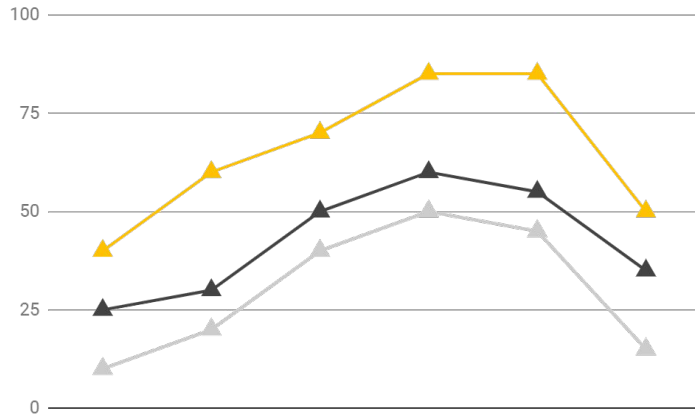
MÉTODOS CUALITATIVOS

El pasado no proporciona una información directa sobre el fenómeno considerado

MÉTODOS CUANTITATIVOS

Se extrae toda la información posible contenida en los datos y, en base al patrón de conducta seguida en el pasado, realizar predicciones sobre el futuro

Análisis



El análisis de series temporales puede ser visto como la tarea de encontrar patrones en los datos temporales y predecir sus valores. La detección de patrones incluye el análisis de:

- **Tendencias:** cambios sistemáticos no repetitivos (lineales o no lineales) de algún valor sobre el tiempo.
- **Ciclos:** el comportamiento observado durante el tiempo es cíclico.
- **Periodos:** los patrones detectados se repiten durante un período de tiempo determinado, ya sea anual, mensual o diario
- **Anomalías:** la técnica de detección de anomalías elimina mucho de los llamados “falsos positivos”, lo cual ayuda a encontrar patrones

Minería de Datos para Series Temporales

Los métodos basados en minería de datos para series temporales son capaces de manipular series temporales no lineales y no estacionarias. Por lo que resultan más útiles para predecir eventos imprevistos en la serie (como por ejemplo el alza repentina del precio de algún producto en el mercado o la rotura de alguna clase de motor en una fábrica etc.)

La naturaleza de las series temporales hace que su tratamiento se diferencie de los métodos tradicionales de minería de datos. Entre estas características distintivas se encuentran: **alta numerosidad, gran número de dimensiones y una constante actualización de sus datos al transcurrir el tiempo.**

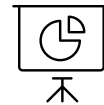
Tareas de la Minería de Datos para Series Temporales



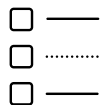
**Representación
e Indexado**



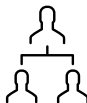
**Emparejamiento
de Subsecuencias**



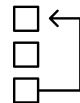
Visualización



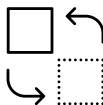
Clasificación



Segmentación



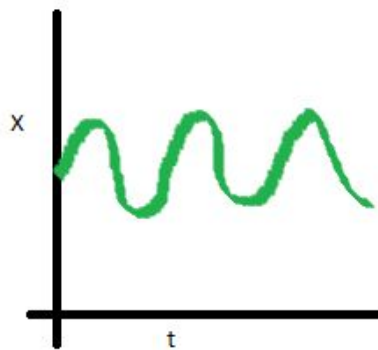
**Descubrimiento
de Patrones y
Conglomerados**



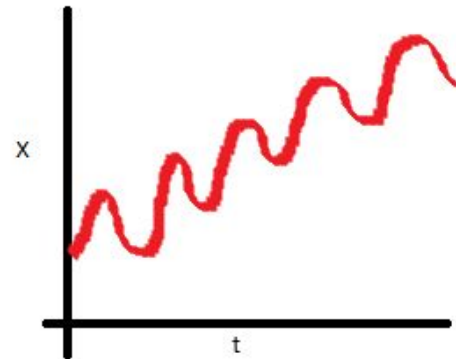
**Medidas de
Similitud**

Tipos de Series de Tiempo

- Estacionarias:
Media y varianza
constantes en el tiempo
- No estacionarias:
Tendencia y/o
variabilidad cambiantes
en el tiempo.



Stationary series



Non-Stationary series

Importancia de trabajar con series estacionarias

Una serie estacionaria es mucho más fácil de predecir.

Si se trabaja con series no estacionarias:

- El principal problema está en la estimación del riesgo.
- Estaremos subestimando el riesgo, porque la distribución tiene colas más largas.
- Estaremos sobrestimando los beneficios.
- El margen de error de las previsiones es mayor

Comprobar Estacionariedad

Por media

Se hace una prueba de hipótesis para descartar tendencia

Por varianza

Se hace una prueba de hipótesis sobre la estacionariedad.
Ej. Dickey-Fuller aumentada

Nota: En caso de no tener una serie de tiempo estacionaria se tienen que hacer las transformaciones necesarias

Modelos de Predicción

- ❖ Modelo Autoregresivo - AR (p)

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \phi_3 x_{t-3} + \dots + \phi_p x_{t-p} + e_t$$

- ❖ Medias Móviles - MA (q)

$$x_t = e_t + \beta_1 e_{t-1} + \dots + \beta_q e_{t-q}$$

- ❖ Modelo ARMA (p,q)

$$x_t = \underbrace{\phi_1 x_{t-1} + \dots + \phi_p x_{t-p}}_{p \text{ términos bajo el modelo AR}} + \underbrace{\beta_1 e_{t-1} + \dots + \beta_q e_{t-q}}_{q \text{ términos bajo el modelo MA}} + c$$

EJEMPLO

