Nombre alumno: Esteban Javier Berumen Nieto

Correo alumno: esteban.berumen@iteso.mx

Instituto tecnológico de Estudios Superiores de Occidente

Repaso para primer examen de la clase Series de Tiempo

Docente: Daniel Nuño, danielnuno@iteso.mx

Otoño 2024

Fecha de entrega: miércoles 27 de noviembre, 2024

El repaso está diseñado en base a los conocimientos teóricos adquiridos en la clase y tomando como referencia el libro Forecasting: Principles and Practice (3rd ed), capítulos 5 a 12.

La entrega es en digital en un archivo de Word. Las preguntas y respuestas deben diferenciarse claramente.

En tus palabras, o subrayando si se presentan opciones, responde las siguientes preguntas:

1. De acuerdo con el libro, ¿cuál es el proceso iterativo de un flujo de pronóstico?
   1. Definición del problema: Entender claramente qué se desea pronosticar.
   2. Recolección de datos: Obtener datos relevantes, precisos y suficientes.
   3. Análisis exploratorio de datos: Identificar patrones, estacionalidad, tendencia, y problemas como datos atípicos.
   4. Selección del modelo: Elegir el modelo más apropiado según los datos.
   5. Ajuste del modelo: Entrenar el modelo utilizando los datos.
   6. Evaluación del modelo: Validar el rendimiento mediante métricas de error.
   7. Uso del modelo para pronosticar: Generar predicciones para el futuro.
   8. Revisión y actualización: Monitorear la precisión del modelo y ajustarlo conforme cambien las condiciones.
2. Los componentes de una serie de tiempo descompuesta son:
   1. Tendencia-ciclo, estacionalidad y residuos.
   2. Media y varianza cero.
   3. Autocorrelación y normalidad.
3. ¿Qué mide la autocorrelación?
   1. Mide la relación lineal entre una variable dependiente y una independiente.
   2. Mide la relación lineal entre valores rezagados de una serie de tiempo.
   3. Mide la tendencia de la serie de tiempo.
   4. Mide la normalidad de la serie de tiempo.
4. ¿Para qué sirve la prueba Ljung–Box? ¿qué significa aceptar la hipótesis nula y cuando aceptas la hipótesis nula?
   1. La prueba Ljung–Box evalúa si los residuos de un modelo están correlacionados entre sí.
   2. Hipótesis nula: Los residuos no están autocorrelacionados (independencia).
   3. Aceptar: Significa que el modelo captura adecuadamente la autocorrelación, y los residuos son ruido blanco.
   4. Rechazar: Indica que persiste autocorrelación en los residuos, sugiriendo que el modelo necesita mejoras.
5. ¿Para qué sirve la prueba Shapiro-Wilk? ¿qué significa aceptar la hipótesis nula y cuando aceptas la hipótesis nula?
6. Esta prueba evalúa si los datos siguen una distribución normal.
7. Hipótesis nula (H₀): Los datos tienen una distribución normal.
8. Aceptar H₀: Indica que los datos no presentan desviaciones significativas respecto a la normalidad.
9. Rechazar H₀: Sugiere que los datos no son normales.
10. ¿Para qué sirve la prueba Breusch–Pagan o White? ¿qué significa aceptar la hipótesis nula y cuando aceptas la hipótesis nula?
11. Ambas pruebas evalúan la presencia de heterocedasticidad en los residuos.
12. Hipótesis nula (H₀): No hay heterocedasticidad; la varianza de los errores es constante.
13. Aceptar H₀: Indica que los residuos son homocedásticos (cumplen el supuesto de varianza constante).
14. Rechazar H₀: Sugiere heterocedasticidad, lo que puede sesgar las estimaciones del modelo.
15. ¿Para qué sirve la prueba Dickey-Fuller o KPSS? ¿qué significa aceptar la hipótesis nula y cuando aceptas la hipótesis nula?
16. Ambas pruebas determinan la estacionariedad de una serie de tiempo.
17. Dickey-Fuller:
18. H₀: La serie tiene una raíz unitaria (no estacionaria).
19. Aceptar H₀: La serie no es estacionaria.
20. Rechazar H₀: La serie es estacionaria.
21. KPSS:
22. H₀: La serie es estacionaria.
23. Aceptar H₀: La serie es estacionaria.
24. Rechazar H₀: La serie no es estacionaria.
25. ¿Cuáles son las características de los errores cuando un modelo se ajusta correctamente a los datos?

Los errores son independientes (sin autocorrelación).

Los errores tienen media cero.

Los errores son homocedásticos (varianza constante).

Los errores son normalmente distribuidos.

1. ¿Cómo estabilizas la media?

Aplicando una diferenciación

1. ¿Cómo estabilizas la varianza?

Aplicando una transformación matemática como puede ser boxcox

1. ¿cuál es la diferencia entre la regresión lineal simple y la regresión lineal multiple?

En la simple solo usamos una variable predictora mientras que en la multiple usamos 2 o mas variables predictoras

1. Dado los datos (X, y) escribe la función de costo para la regresión lineal y explica cómo el método de mínimos cuadrados la minimiza.

La función de costo en regresión lineal es el error cuadrático medio (MSE):

J(β0,β1,...,βn)=1n∑i=1n(yi−y^i)2

Método de mínimos cuadrados: Encuentra los valores de los parámetros ($\beta$) que minimizan $J$ resolviendo el sistema de ecuaciones:

∂J/∂βj=0 para cada j

1. Supón que tu regresión lineal quieres reducir la cantidad de variables predictoras que estas usando. Explica que harías y por que

Eliminar variables irrelevantes y fijarme en la correlación de las variables con mi obetivo para, usando distintos métodos para ver cuales son las variables mas relevantes para realizar las predicciones del modelo

1. Describe para que tipo de series de tiempo (que características tienen los datos) usarías los modelos de Suavización Exponencial (ES)

Datos con tendencia moderada y/o estacionalidad

Patrones históricos que se espera continúen en el futuro.

Series sin variaciones abruptas ni comportamiento caótico.

1. Supongamos que utiliza el método Holt-Winters (A,A) para predecir ventas con tendencia y estacionalidad. Tras ajustar el modelo, observa que tus pronostico están sistemáticamente sesgadas durante determinadas estaciones. ¿Qué medidas tomaría para diagnosticar y mejorar su modelo?

**Diagnóstico**:

1. Analizar los residuos para detectar patrones estacionales persistentes.
2. Realizar pruebas de autocorrelación (ACF, Ljung-Box) para evaluar si hay correlación no capturada.

**Mejora**:

* Revisar si la periodicidad de la estacionalidad está bien especificada.
* Probar modelos con diferentes configuraciones del método (e.g., multiplicativo).
* Ajustar manualmente los parámetros iniciales o usar optimización numérica para afinarlos.
* Considerar un modelo diferente, como SARIMA, si los patrones estacionales son más complejos.

1. Escribe la fórmula matemática de la suavización exponencial simple y explica el papel del parámetro de suavización . ¿Cómo afecta el modelo a los cambios recientes en los datos dado por ?

*y*^​*t*​=*αyt*​+(1−*α*)*y*^​*t*−1​,

Valores altos de alpha: Dan más peso a los datos recientes, respondiendo rápidamente a cambios.

Valores bajos de alpha: Dan más peso al historial pasado, produciendo una serie más suave pero menos reactiva.

1. Describe para que tipo de series de tiempo (que características tienen los datos) usarías los modelos ARIMA

No estacionarias que pueden transformarse en estacionarias mediante diferenciación.

Sin patrones estacionales claros (usar SARIMA para series con estacionalidad).

Series donde la relación entre los valores pasados y futuros se captura bien con rezagos.

1. Explica los componentes de un modelo ARIMA(p, d, q). ¿Qué significan los parámetros y cómo se determinan?

p (autoregresivo): Número de rezagos de la variable dependiente (y\_t) que se incluyen en el modelo. Determinado por la gráfica PACF.

d (diferenciación): Veces que se diferencia la serie para hacerla estacionaria. Determinado por pruebas de estacionariedad (ADF, KPSS).

q (media móvil): Número de rezagos de los errores (epsilon\_t) incluidos en el modelo. Determinado por la gráfica ACF.

1. El parámetro d en ARIMA significa el «orden de diferenciación». ¿Por qué es necesaria la diferenciación y cómo se determina el número óptimo de diferencias que hay que aplicar a una serie temporal?

Razón: La diferenciación elimina tendencia y convierte una serie no estacionaria en estacionaria, una condición necesaria para ARIMA.

Determinar d:

Analizar gráficos de la serie y su diferencia (visualizar si la varianza y media son constantes).

Aplicar pruebas de estacionariedad (e.g., ADF, KPSS).

Incrementar d hasta que la serie sea estacionaria (sin sobre-diferenciar).

1. ¿Cómo agregas el componente estacional a un modelo ARIMA?

Usando un modelo SARIMA (Seasonal ARIMA), con componentes adicionales para estacionalidad:

SARIMA(p,d,q)(P,D,Q,s)

donde:

P: Orden autoregresivo estacional.

D: Orden de diferenciación estacional.

Q: Orden de media móvil estacional.

s: Periodicidad estacional (e.g., 12 para datos mensuales con un ciclo anual).

Se determinan observando ACF/PACF en los rezagos estacionales y evaluando estacionariedad en las diferencias estacionales.

1. Una regresión dinámica utiliza una regresión lineal y ajusta el modelo ARIMA sobre los \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Residuos

1. Si tienes una estructura jerarquía de series de tiempo con diferentes niveles, como el índice S&P 500, y quieres pronosticar el futuro valor del índice ¿Es mejor pronosticar el nivel más alto (el valor del índice) o el nivel más bajo (cada acción)?

Nivel más bajo (acciones individuales): Es más detallado y puede capturar mejor la dinámica individual. Luego, los resultados se agregan para obtener el pronóstico del índice. Es preferible cuando los datos son confiables y representativos.

Nivel más alto (índice directamente): Es más simple y rápido. Es útil si el interés principal está en el índice y no se necesita información granular.

1. Describe la idea principal del vector autoregresión.

El VAR es un modelo estadístico multivariante que extiende el modelo autoregresivo para capturar relaciones dinámicas entre múltiples series de tiempo.

Cada variable en el sistema se modela como una función lineal de los valores pasados de sí misma y de las otras variables.

Es útil para analizar y pronosticar sistemas donde las variables están interrelacionadas, como economía, clima o mercados financieros.

1. Describe la idea principal de las redes neuronales.

Las redes neuronales son modelos de aprendizaje automático inspirados en el cerebro humano que capturan relaciones complejas en los datos.

Componentes básicos: Neuronas organizadas en capas (entrada, ocultas y salida).

Usan funciones de activación para transformar datos y aprender patrones no lineales.

Aplican optimización para minimizar una función de pérdida ajustando los pesos de las conexiones.

Son útiles para problemas de predicción, clasificación y series temporales, especialmente cuando las relaciones entre las variables son complejas y no lineales.

1. Describe la idea principal de Bootstrapping and bagging.

Bootstrapping y Bagging son técnicas de muestreo y combinación usadas en aprendizaje automático para mejorar la precisión y estabilidad de los modelos.

Bootstrapping: Crea múltiples subconjuntos de datos mediante muestreo aleatorio con reemplazo del conjunto original.

Bagging (Bootstrap Aggregating): Entrena múltiples modelos independientes (usando los subconjuntos generados) y combina sus resultados (por ejemplo, promedio para regresión o voto mayoritario para clasificación).