

**LABORATORIO DE INTELIGENCIA COMPUTACIONAL
FORECAST DE DEMANDA Y DISEÑO DE EXPERIMENTOS
29 DE MARZO 2020**

1. Trabajo Previo

Descargar Python 3.7 desde el sitio Anaconda.org ➔ Download Anaconda

Una vez descargado Anaconda, instalarlo.

Al terminar, desde línea de comandos escribir jupyter notebook y presionar enter. Esto abrirá una sesión web mostrando las carpetas del computador local.

Descargar de U-Cursos el notebook llamado Serie Temporal.ipynb. Guardarlo en una carpeta del computador local.

Desde la sesión web de Jupyter navegar hasta encontrar la carpeta donde se guardó el notebook. Dar click en el notebook para estudiar los conceptos análisis asociados a una predicción de series de tiempo.

Suponga que es el VP de Business Analytics (Inteligencia Computacional) de una compañía de Consumo Masivo (CPG) que importa vinos desde Australia para comercializarlos a través de canales, y que también produce textiles basados en fibra sintética. El CEO de la compañía le ha encargado 2 desafíos de negocio que permita mejorar las ventas y margen en la venta de vinos australianos y baje los costos de producción de los artículos textiles para obtener mayor margen, utilizando Inteligencia Computacional que permita tomar decisiones basada en los datos más allá del juicio de expertos.

2. Forecast de ventas de vino para pronosticar las ventas de vinos australianos usando Python

Utilizar los datos del siguiente archivo:

<http://www.forecastingbook.com/mooc/AustralianWines.csv?attredirects=0&d=1> "

Se necesita:

- 2.1 Realizar un análisis de las ventas mensuales de vinos australianos. Los datos corresponden a seis tipos de vinos (tinto, rosa, blanco dulce, blanco seco, espumoso y fortificado) para el período 1980-1994. Las unidades son miles de litros. Se le contrata para obtener pronósticos a corto plazo (2 meses por adelantado) para cada una de las seis series y esta tarea se repetirá cada mes a medida que lleguen nuevos datos.
- 2.2 Evaluar el desempeño de pronósticos examinando el desempeño del conjunto de entrenamiento y el desempeño del conjunto de validación.

- 2.3 Elegir uno o más tipos de vino.
- 2.4 Calcular pronósticos para los períodos de entrenamiento y validación.
- 2.5 Trazar los valores reales y pronosticados para los períodos de entrenamiento y validación.
- 2.6 Calcular medidas de rendimiento (por separado) para los períodos de entrenamiento y validación.
- 2.7 Crear un histograma de los errores de pronóstico del período de validación. Examinar la forma en términos de magnitudes de error. Establezcamos nuestro nivel de confianza deseado al 80%.
- 2.8 Calcular los deciles primero y noveno de la distribución de errores de validación. Combinar los dos deciles con el pronóstico puntual para obtener un intervalo de predicción del 80%.
- 2.9 Usar la diferenciación para eliminar la tendencia y / o la estacionalidad de cada serie para que luego podamos usar un promedio móvil para pronosticar las ventas futuras.
- 2.10 Escribir sobre las operaciones de diferenciación que aplicó a cada una de las seis series, explicando por qué eligió esas operaciones. Incluir gráficos puede ayudar a respaldar sus argumentos.
- 2.11 Predicción de ventas de vinos utilizando un MA: Use un promedio móvil con un ancho de ventana de 12 (MA (12)) para pronosticar las ventas con doble diferencia de las ventas de vino fortificado en febrero de 1995.
- 2.12 Finalmente, ajuste el pronóstico "sin diferenciar" dos veces, para incluir de nuevo la tendencia y la estacionalidad. Reportar su pronóstico final para las ventas de vinos fortificados en febrero de 1995 utilizando este método. Discuta cualquier desafío o pregunta que haya encontrado.

3. Diseño de Experimentos (DoE) para mejorar los costos de producción de una empresa de manufactura que utiliza fibra sintética

La siguiente tabla presenta los datos de resistencia a la tensión de 5 tipos diferentes de fibra sintética. Mientras mayor porcentaje de algodón, mayor costo de producción. Se ha realizado mediciones (observaciones) sobre 5 muestras para cada tipo de fibra sintética.

Cotton Percentage	Observations				
	1	2	3	4	5
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

A través de análisis de datos sobre un experimento estadísticamente diseñado (utilizando Design of Experiments – DoE) concluya si existe diferencia significativa en la resistencia a la tensión entre los 5 tipos de fibra evaluados. Realice el análisis completo utilizando Python.

De NO haber diferencia significativa... ¿Tendrá algún sentido producir la fibra con mayor porcentaje de algodón? O dicho de otra manera... ¿Tendrá algún sentido producir una fibra más

costosa si su resistencia a la tensión es la misma que las otras? Justifique elegantemente su respuesta.... De allí su calificación.

Más información en el siguiente link:

<http://www.um.edu.ar/math/montgomery.pdf#page=3&zoom=100,0,0>

Entregar hasta el jueves 30 de abril de 2020 por U-Cursos.

PAUTA INFORME

El informe desarrollado debe contener al menos los siguientes puntos:

1. **Índice de contenidos**
2. **Introducción:** Se debe dar una descripción general del problema de negocio a resolver, una breve descripción del informe y una clara visualización de objetivos, tanto generales como específicos. No debe ser una copia del enunciado.
3. **Análisis estadístico y exploratorio de los datos:** En este punto deben calcular las principales métricas estadísticas para una mejor descripción de los datos, por ejemplo estadísticas descriptivas, tablas de frecuencia, histogramas, kurtosis, asimetría de aplicar.
4. **Algoritmos de Machine Learning:** En esta sección se debe comparar en términos de medidas de bondad asociadas a las tareas analíticas. Se debe realizar una pequeña sensibilización de los parámetros de ajuste que definen a los modelos respecto a la medida de performance.
5. **Determinar el mejor resultado basado en los datos y análisis ejecutado.**
6. **Políticas comerciales y toma de decisiones para el negocio.**
7. **Conclusiones y discusiones.**
8. **Anexos (no en exceso ya que no sumarán en la evaluación).**