



Taller de Aprendizaje Automático Obligatorio

**Diploma de Especialización de
Analítica Big Data, ORT**

Mayo de 2019

Aspectos Formales

- **Entrega (45 puntos):**
 - Modalidad: grupos de a 2 o 3 personas.
 - Entregable: Word con script en R como anexo o archivo R markdown/notebook.

Presentación del caso de negocio

Rest-on es una guía turística que está armando en su web una sección de restaurantes, que sirva para asesorar a sus clientes en las características y calidad de los restaurantes.

A estos efectos, los contratan a ustedes para apoyarlos en los algoritmos de soporte de su web, con dos objetivos:

- i. Poder recomendarles a sus clientes restaurantes similares a los que han consultado
- ii. Rankear a los restaurantes, en términos relativos a los restaurantes comparables.

A estos efectos, cuenta con una base de datos de restaurantes, a la cual se puede acceder en el siguiente link:

<https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data>

Pauta del obligatorio: Entendimiento del Caso de Negocio y Planificación del Trabajo

1. Discuta el desafío planteado en términos de la importancia que tiene para una empresa cualquiera y para Rest-on en particular.
2. Identifique conceptual y técnicamente qué tipo de modelos serían necesarios para atender este problema (no el algoritmo específico, sino el tipo de técnica).
3. Plantee hipótesis respecto a qué dimensiones de análisis (a nivel conceptual y macro) y variables pueden ser relevantes para lograr el objetivo.
4. Identifique, en base a lo anterior, qué tipo de datos precisaría obtener para llevar a cabo este trabajo.
5. Arme un plan de trabajo, identificando las diferentes etapas y actividades para lograr el objetivo.
6. ¿Cómo será utilizado el resultado del trabajo por el cliente?
7. En base a lo anterior, ¿Qué áreas se deberían involucrar en el proyecto y qué rol cumplirían?

Pauta del obligatorio: Extracción, Transformación y Carga de Datos

1. Diseñe la estructura de la tabla de datos analítica: ¿qué va a ser cada fila? ¿cuáles van a ser las columnas?
2. ¿Cómo se construye cada una de las variables en la tabla de datos? ¿De qué tablas surge y qué transformaciones de datos son necesarias?
3. Construya la tabla de datos analíticas en base a la especificación anterior.

Pauta del obligatorio: Exploración y análisis descriptivo

1. Realice un análisis descriptivo de la tabla de datos.
2. ¿Descarta a priori alguna variable a partir de este análisis?
3. ¿Puede sospechar qué variables serán útiles para el problema identificado?
4. ¿Es necesario tomar alguna decisión respecto al universo o a la exclusión de valores atípicos?

Pauta del obligatorio: Modelado y Evaluación

1. ¿Qué variables pre-seleccionaría en base al análisis funcional y estadístico?
2. Estime al menos dos modelos (usando diferentes algoritmos y/o variables), compárelos (en base a criterios estadísticos y funcionales) y seleccione el que entienda más adecuado.
3. Caracterice y describa los resultados a los cuales se arribaron en el caso anterior, desde el punto de vista del negocio.

Pauta del obligatorio: Implementación

1. ¿Cómo entiende que el modelo podría ser utilizado por la web para cumplir el objetivo funcional?
2. En base a lo anterior, ¿cómo debería ser la implementación de este modelo (estimación resultados en tiempo real, procesamiento batch, etc)?
3. ¿Cómo se podrían armar, usando los resultados de este modelo, los rankings de restaurantes “en términos relativos” o “específicos a un grupo”? Si bien no es necesario realizar los desarrollos, identifique el proceso, actividades, variables a utilizar, etc y describa con el mayor detalle posible.
4. ¿Qué ventaja tendría esto frente a usar un ranking general?
5. ¿Cómo se podría evaluar el resultado (éxito) del trabajo?



Taller de Aprendizaje Automático Obligatorio

**Diploma de Especialización de
Analítica Big Data, ORT**

Mayo de 2019