**Caso de estudio: Empresa de Agua Potable**

Una empresa de agua potable brinda el servicio de agua potable a miles de abonados (clientes). Por cada medidor de agua instalado en una casa se ha creado una cuenta. Una cuenta de un abonado se encuentra ubicada en un ciclo (Parroquias) y está categorizada según el tipo de vivienda o comercio. Cada mes emite una carta de pago (factura), la cual debe ser cancelada; sin embargo, no todos los clientes son puntuales en sus pagos. Estas cartas impagas forman parte de una cartera de cobro.

La empresa requiere que se elabore un sistema de soporte de decisiones (DSS), con interfaz web y estilo dashboard, que incluya estadística básica (gráficos y cuadros) y análisis de datos avanzado (técnicas de machine learning y/o minería de datos). A continuación, se lista la información mínima que debe presentar la aplicación:

1. Información que debe permitir ser filtrada por: ciclo o parroquia, categoría, año, emisión (año-mes), abonado deudores y no deudores **(1 ptos):**
   1. Cantidad de abonados
   2. Total (valor a pagar por el servicio) emitido
   3. Total pagado o recaudado
   4. Total impago (deuda)
   5. Gráficos de total recaudado por:
      1. Años
      2. Ciclo
      3. Categoría
      4. emisión (año-mes)
   6. Gráficos de deuda por:
      1. Años
      2. Ciclo
      3. Categoría
      4. emisión (año-mes)
2. Seleccionar una cuenta de un abonado, colocar filtros: todo, por año, o por emisión (año-mes) **(1 ptos):**
   1. Gráfico y tabla de pagos y deudas
   2. Cantidad de mese que adeuda
   3. Total (valor a pagar por el servicio) emitido
   4. Total pagado o recaudado
   5. Total de deuda
   6. Año donde comienza la deuda
3. Análisis avanzado:
   1. Aplicar análisis de correlación seleccionando todas las variables numéricas de interés: cantidad de meses del todo el servicio recibido, cantidad de meses que adeuda, cantidad de meses pagados, total (valor a pagar) emitido, total pagado, total de deuda, ciclo, categoría. Determinar matriz de coeficientes de correlación (identificar variables y valores significantes que tienden a |1|). **(0.75 ptos)**
   2. Regresión lineal (relación entre dos variables numéricas) entre cantidad de mese que adeuda versus total de la deuda. Predecir la deuda en 3, 6 y 12 meses posteriores. Aplicar filtros por ciclo, categoría y abonado. Determinar la eficiencia del modelo. **(0.75 ptos)**
   3. Serie temporal (eje x: año, eje y: total de la deuda) Predecir la deuda en 3, 6 y 12 meses posteriores. Aplicar filtros por ciclo, categoría y abonado. Determinar la eficiencia del modelo. **(0.75 ptos)**
   4. Mediante técnicas de clústering, encontrar grupos de abonados que tienen características similares y describir cuáles son esas características que definen a cada grupo. **(0.75 ptos)**
   5. Aplicar bosques aleatorios y redes neuronales como modelo s de clasificación que ayuden a determinar si un abonado es pagador o no). Para probar el modelo ingresar datos de un abonado: cantidad de meses del todo el servicio recibido, cantidad de meses que adeuda, cantidad de meses pagados, total (valor a pagar) emitido, total pagado, total de deuda, ciclo, categoría. **(1 ptos)**