A continuación, encontrará lo que compartimos con usted en los diferentes momentos del parcial. Caso inicial – antes del parcial, Etapa 1 durante la primera sesión del parcial. Etapa 2 durante la segunda sesión del parcial.

**Caso Inicial Clínica de los Alpes**

La Clínica de los Alpes es una IPS colombiana con varios años de operación en la capital. Esta institución posee varias unidades de cuidado especializado para pacientes con diferentes tipos de patologías y necesidades de atención.

Sin embargo, al igual que muchas entidades de las mismas características, uno de los mayores problemas que enfrenta la Clínica es el de los reingresos médicos. Un reingreso médico ocurre cuando un paciente que se ha dado de alta debe ser nuevamente admitido por complicaciones relacionadas con su ingreso original en plazos de 30 días, 90 días o hasta un año, de acuerdo con las definiciones de la OMS.

Los reingresos son problemáticos porque pueden ser un signo de complicaciones o empeoramiento de la condición médica del paciente, lo que pone en riesgo su salud y bienestar. Esto puede generar frustración y ansiedad por parte del paciente al igual que aumentar el riesgo de contraer infecciones dentro de la clínica. De igual manera, aumentan la carga financiera sobre la clínica, añadiendo costos acumulativos de cuidado, bloqueando el uso eficiente de recursos para atender a otros pacientes, generando sobrecostos para las facturas médicas que deben pagar los pacientes, causando impactos físicos y emocionales; y en últimas ocasionando pérdida de confianza en la institución de salud.

**Etapa 1 [50%]. En esta etapa se entrega el caso que se describe a continuación, los datos y el diccionario de datos.**

**INCLUIR ESTE PÁRRAFO EN EL CUADERNO QUE ENTREGA COMO SOLUCIÓN AL PARCIAL.**

“Al entregar la solución de este parcial, yo, **[nombre y apellido]** con código **[#]** me comprometo a no conversar durante el desarrollo de este examen con ninguna persona que no sea el profesor del curso, sobre aspectos relacionados con el parcial; tampoco utilizaré algún medio de comunicación por voz, texto o intercambio de archivos, para consultar o compartir con otros, información sobre el tema del parcial. Ni usaré alguna IA para resolver las preguntas del parcial. Soy consciente y acepto las consecuencias que acarreará para mi desempeño académico cometer fraude en este parcial”.

La Clínica de los Alpes es una IPS colombiana con varios años de operación en la capital. Esta institución posee varias unidades de cuidado especializado para pacientes con diferentes tipos de patologías y necesidades de atención.

Sin embargo, al igual que muchas entidades de las mismas características, uno de los mayores problemas que enfrenta la Clínica es el de los reingresos médicos. Un reingreso médico ocurre cuando un paciente que se ha dado de alta debe ser nuevamente admitido por complicaciones relacionadas con su ingreso original en plazos de 30 días, 90 días o hasta un año, de acuerdo con las definiciones de la OMS.

Los reingresos son problemáticos porque pueden ser un signo de complicaciones o empeoramiento de la condición médica del paciente, lo que pone en riesgo su salud y bienestar. Esto puede generar frustración y ansiedad por parte del paciente al igual que aumentar el riesgo de contraer infecciones dentro de la clínica. De igual manera, aumentan la carga financiera sobre la clínica, añadiendo costos acumulativos de cuidado, bloqueando el uso eficiente de recursos para atender a otros pacientes, generando sobrecostos para las facturas médicas que deben pagar los pacientes, causando impactos físicos y emocionales; y en últimas ocasionando pérdida de confianza en la institución de salud.

Por esta razón la Clínica de los Alpes ha financiado un proyecto de analítica basada en aprendizaje automático con el propósito de caracterizar los pacientes con reingresos, en particular los que han sido hospitalizados por diabetes.

Considerando el caso asignado, los datos y el diccionario suministrado, proponga un proyecto con aprendizaje automático para alcanzar uno de los objetivos planteados por la empresa. Resuelva las siguientes preguntas e incluya sus respuestas y análisis en el cuaderno a entregar.

1. **[10%] Entendimiento del negocio y enfoque analítico**
   1. **[3%]** Describa el objetivo de negocio que tiene la clínica de los Alpes donde es posible aplicar aprendizaje automático. Justifique la razón por la cual el uso de aprendizaje automático es una opción para ellos.

El objetivo de negocio es caracterizar los pacientes con reingresos, en particular

Identificar el problema de negocio y hacer un match con el proceso de aprendizaje. Como quiero caracterizar, necesito realizar una agrupación.

Regresión predicción de un numero

Clasificacion predicción de una clase

Traten de ser breves…

* 1. **[7%]** Describa y justifique el enfoque analítico propuesto para abordar el objetivo de negocio identificado en el punto 1a. Incluya los tipos y tareas de aprendizaje, al igual que las técnicas o algoritmos de aprendizaje automático que propone utilizar.

TAREA: Tipo no supervisado y una tarea de agrupación. Se empleara un K-means (permite facilitar el análisis para poder hacer la caracterización, analizado en prototipos)…

Uno es suficiente, cumplan el proceso de ML!!!

1. **[20%]** Analice el resultado del entendimiento de los datos compartidos según la metodología CRISP-ML. Recuerde incluir el análisis relacionado con el perfilamiento y la calidad de los datos.

Perfilamiento de los datos (exploración), se tiene que conocer los datos.

Exploración, limpieza y preparación

SE DEBE ANALIZAR EL RESULTADO DE CADA ETAPA

CADA ALGORITMO tiene una mejor forma de preparar los datos para el modelo

SECUENCIA A SEGUIR

A nivel de evaluación, siempre hay que sacar las métricas sobre entrenamiento y test, si son las métricas adecuadas. Hay que balancear,

1. **[20%]** Explique y justifique las tareas de transformación realizadas sobre los datos para aplicar luego el enfoque analítico propuesto. En este punto debe implementar las transformaciones propuestas y mostrar el resultado obtenido luego de ejecutar el código desarrollado.

**Condiciones de entrega**:

* La entrega se hace por BloqueNeón y corresponde al cuaderno **ejecutado** que construyeron para este enunciado en formatos “ipynb” y html o pdf (en caso de tener dificultad para generar el html). En JupyterLab el formato html se crea con la opción “File/Export As”.
* Recuerde incluir el código de honor y la respuesta a cada una de las preguntas planteadas por la institución, en la primera parte del cuaderno. En esta parte recuerde referenciar las celdas del cuaderno que soportan su descripción.

El nombre de los archivos a entregar debe ser: loginUniandes\_etapa1\_parcial1.

**Etapa 2 [50%] En esta etapa se entrega el caso que se describe a continuación, los datos que no son necesariamente los mismos de la etapa1, el diccionario de datos y un cuaderno con el desarrollo de un modelo analítico.**

**IDENTFICAR 2 ERRORES Y COMO LO IMPLDENTARONÑ.**

**INCLUIR ESTE PÁRRAFO EN EL L CUADERNO QUE ENTREGA COMO SOLUCIÓN AL PARCIAL.**

“Al entregar la solución de este parcial, yo, **[nombre y apellido]** con código **[#]** me comprometo a no conversar durante el desarrollo de este examen con ninguna persona que no sea el profesor del curso, sobre aspectos relacionados con el parcial; tampoco utilizaré algún medio de comunicación por voz, texto o intercambio de archivos, para consultar o compartir con otros, información sobre el tema del parcial. Ni usaré alguna IA para resolver las preguntas del parcial. Soy consciente y acepto las consecuencias que acarreará para mi desempeño académico cometer fraude en este parcial”.

La Clínica de los Alpes es una IPS colombiana con varios años de operación en la capital. Esta institución posee varias unidades de cuidado especializado para pacientes con diferentes tipos de patologías y necesidades de atención.

Sin embargo, al igual que muchas entidades de las mismas características, uno de los mayores problemas que enfrenta la Clínica es el de los reingresos médicos. Un reingreso médico ocurre cuando un paciente que se ha dado de alta debe ser nuevamente admitido por complicaciones relacionadas con su ingreso original en plazos de 30 días, 90 días o hasta un año, de acuerdo con las definiciones de la OMS.

Los reingresos son problemáticos porque pueden ser un signo de complicaciones o empeoramiento de la condición médica del paciente, lo que pone en riesgo su salud y bienestar. Esto puede generar frustración y ansiedad por parte del paciente al igual que aumentar el riesgo de contraer infecciones dentro de la clínica. De igual manera, aumentan la carga financiera sobre la clínica, añadiendo costos acumulativos de cuidado, bloqueando el uso eficiente de recursos para atender a otros pacientes, generando sobrecostos para las facturas médicas que deben pagar los pacientes, causando impactos físicos y emocionales; y en últimas ocasionando pérdida de confianza en la institución de salud.

Por esta razón la Clínica de los Alpes ha financiado un proyecto de analítica basada en aprendizaje automático con el propósito de detectar cuando un paciente que está a punto de ser dado de alta tiene un riesgo considerable de reingreso y por lo tanto requiere atención adicional. En la fase inicial del desarrollo de una herramienta de detección de riesgos de reingresos, la clínica ha decidido concentrarse en los pacientes con diabetes, una de las condiciones de salud de pacientes que se ve asociada con un mayor riesgo a ser reingresado por complicaciones. El equipo de desarrollo ha decidido crear un clasificador usando un conjunto de datos de pacientes con esta enfermedad.

Esta primera versión ha sido desplegada en una aplicación interna para la clínica, que analiza los pacientes con diabetes internados, de tal manera que antes de ser dados de alta, se pueda determinar si poseen un alto riesgo de reingreso y por lo tanto requieren tiempo adicional de tratamiento y/o observación.

La impresión de la junta directiva de la institución es que esta aplicación tendrá un gran éxito, debido a que el equipo de desarrollo obtuvo una exactitud (*accuracy*) de cerca de 68% sobre los datos usados. Ante la percepción de la efectividad de la solución actual, los directivos han expresado que cualquier mejora sobre la solución debe ser probada cautelosamente, con el cuidado de no afectar innecesariamente el bienestar de los pacientes, ni generar malestar entre el personal de salud que está adoptando la herramienta aún con cierto recelo.

Resolver los siguientes puntos tomando como base el diccionario y los datos compartidos, al igual que el modelo de clasificación entregado en el cuaderno anexo.

1. **[14%]** Identifique y justifique los errores y malas prácticas, según los criterios dados a continuación, del código compartido para crear el modelo de clasificación. Indique el problema y la celda o celdas a nivel de código donde se presenta.
   1. Ingeniería de características
      1. Selección de variables
      2. Preparación de variables
   2. Desarrollo del modelo
   3. Evaluación del modelo

TRANSFORMADORES SE AJUSTAN AL TRAIN Y SE PROCESA EL

SOLO SE ENTRENA CON EL TRAIN

Desarrollo del modelo, no se buscan los hyperparametros

Solución donde se corrigen dos cosas malas

Aprendan a identificar, a nivel de caracteriscticas,

Vean el notebook y a partir de estos tres puntos que contestarían

1. **[8%]** Plantee dos estrategias para corregir errores y malas prácticas identificados en el punto anterior, a nivel de ingeniería de característica y de evaluación del modelo.
2. **[20%]** Implemente las estrategias planeadas en el punto anterior y compare los resultados obtenidos con este segundo modelo para sugerir a la empresa el modelo que debe utilizar. Justifique su recomendación.
3. **[8%]** Explique a las directivas de la clínica de los Alpes cómo el resultado del modelo de clasificación que construyó aporta al objetivo que tiene el negocio de identificar los pacientes hospitalizados por diabetes con un alto riesgo de reingreso.

**Condiciones de entrega**:

* La entrega se hace por BloqueNeón y corresponde al cuaderno **ejecutado** que construyeron para este enunciado en formatos “ipynb” y html o pdf (en caso de tener dificultad para generar el html). En JupyterLab el formato html se crea con la opción “File/Export As”.
* Recuerde incluir el código de honor y la respuesta a cada una de las preguntas planteadas por la institución, en la primera parte del cuaderno. En esta parte recuerde referenciar las celdas del cuaderno que soportan su descripción.
* El nombre de los archivos a entregar debe ser: loginUniandes\_etapa2\_parcial1.

**¡Éxitos!**