

DIPLOMADO Ciencia de Datos**Examen 2**

Nombre: Rodríguez Fitta José Emanuel **Fecha:** 06/03/2022
Apellido Paterno Apellido Materno Nombre(s)

Calificación: _____

Objetivo: *El participante identificará las ventajas y desventajas de las metodologías KDD, Agile design y CRISP para un caso de estudio proporcionado por el participante.*

Instrucciones:

1. Llenar cada uno de los recuadros siguientes explicando el caso de estudio a tratar, así como sus objetivos y los objetivos de minería de datos.
2. Llenar cada uno de los recuadros explicando las ventajas y desventajas al utilizar cada una de las metodologías aplicadas al caso de estudio elegido por el participante.

Explicar el caso de estudio:

Valor 2.5 puntos

Una institución financiera que se dedica a dar tarjetas de crédito ha detectado que cierto porcentaje de sus clientes cometen fraude y los ha registrado en una base de datos donde se incluyen características, como la fecha de activación el balance en su tarjeta, etc. Los fraudes hechos por estos clientes traen consigo pérdidas económicas considerables para la empresa por lo que se busca hacer un algoritmo que pueda predecir a los probables clientes que puedan cometer fraude, se tendría en la mira a dichos clientes y en caso de ser necesario tomar medidas que eviten el fraude y con ello las pérdidas económicas.

Mencionar las ventajas y desventajas utilizando la metodología KDD en el caso de estudio:

Valor 2.5 puntos

	Ventajas	Desventajas
KDD	Es una metodología que en lo general se adecua a lo que se busca, dado que se requiere encontrar patrones dentro de los datos que nos digan que condiciones deben cumplir aquellas personas con tendencia a hacer fraude. Una vez encontrados dichos patrones, se aplicará este conocimiento para evitar que nuevos clientes con las mismas tendencias cometan fraude.	Es un modelo que, si bien en lo general se adecua a los propósitos del proyecto, este, no especifica puntualmente lo que se debe hacer, lo cual podría ocasionar que, al querer replicar todo el proceso en el futuro, esto no se lleve a cabo de manera adecuada. Otra desventaja es que la metodología KDD no contempla la gestión del proyecto, esto puede traer como consecuencia que la terminación del proyecto se alargue.

Mencionar las ventajas y desventajas utilizando la metodología Agile design en el caso de estudio:

Valor 2.5 puntos

	Ventajas	Desventajas
Agile design	<p>Algunas ventajas que se encuentran dentro de la metodología agile es por ejemplo que es un proceso iterativo, dado que las personas que cometen fraude se reinventan constantemente, es importante estar reinventando el modelo junto con ello. Otra ventaja es sin duda el control de versiones puesto que, en caso de cometerse algún error durante algún paso, se puede regresar a una versión anterior, además de que de llevarse a cabo la implementación de un modelo, si este es menos eficiente en la realidad que una versión anterior puede regresarse a esta sin problemas. Otra gran ventaja es que se la da prioridad a lo que los datos dicen, esto es importante por que en ellos se encuentran los patrones que siguen aquellas personas que cometen fraude y sin duda otra gran ventaja es que se basa en llevar a cabo experimentos, esto es importante puesto que puede experimentarse creando un modelo considerando solo características del uso de la tarjeta y estudiar sus resultados y después comparar con un modelo creado con base en características demográficas. Además, tener reportado cada paso es importante puesto que se puede replicar lo ya hecho o mejorarlo.</p>	<p>Debido a que no se sigue una estructura tan detallada como en la metodología CRISP-DM, por ejemplo, puede que en algún punto del proceso los equipos se desvíen fácilmente. Otra desventaja es que, dado que se requiere experimentar constantemente y reinventar el modelo a la par de los métodos de fraude, creemos que la metodología requiere un gran nivel de compromiso por parte de los integrantes, incluyendo gran colaboración entre ellos y posiblemente tiempo adicional. Otra desventaja es que no se sabrá en un principio con certeza la duración del proyecto y el costo.</p>

Mencionar las ventajas y desventajas utilizando la metodología CRISP en el caso de estudio:

Valor 2.5 puntos

	Ventajas	Desventajas
--	----------	-------------

CRISP	<p>Esta metodología se adecua al proyecto dado que se requiere un alto grado de comprensión de negocio para entender los patrones que tienen aquellas personas que cometen fraude. Además de que se busca un modelo de predicción para identificar a dichos clientes. Dentro de esta metodología se contempla una gestión de tiempo, lo cual trae como ventaja que se determine en cuanto tiempo se realizaría un despliegue del modelo y con ello tomar algunas medidas provisionales. Otra ventaja es lo bien documentado que se tendría el proyecto para futuras implementaciones o mejoras.</p>	<p>Dado que los datos están desbalanceados, puesto que tan solo el 1% cometen fraude, puede resultar difícil implementar algunos de sus pasos como lo son el modelado. Otra desventaja será el implementar el modelo, esto dado que constantemente las personas que cometen fraude están reinventando las maneras en que llevan a cabo dichas acciones. Además, dado que es una metodología muy detallada y completa, implementar todo el ciclo puede resultar tardado y esto puede resultar en dos cosas: La primera que se cometan muchos fraudes en este tiempo y con ello se tengan pérdidas considerables y la segunda consecuencia puede ser que en el tiempo en que se implementa el modelo, las personas que cometen fraude ya hayan cambiado de métodos.</p>
--------------	---	---

Valor total 10 puntos
