|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Sigla Asignatura | BIY7101 | Nombre de la Asignatura | Minería de datos | Tiempo | 3 módulos |
| Experiencia de Aprendizaje N° 3 | Genera información y conocimiento útil desde las diversas fuentes de datos para permitir la toma de decisiones de acuerdo con las necesidades de la organización. | | | | |
| Nombre del Recurso Didáctico | 3.2.2 Actividad Modelos predictivos | | | | |

1. **Aprendizajes e indicadores de logro**

|  |  |
| --- | --- |
| Aprendizajes (Procedimentales, Actitudinales y conceptuales) | **Indicadores de logro** |
| Evaluar la confiabilidad de los modelos estadísticos en lenguaje Python, de acuerdo al análisis de datos previamente realizado. | Aplica las funciones en el lenguaje definido, que le permiten la obtención de los coeficientes de evaluación para evaluar los modelos estadísticos. |
| Analiza los coeficientes de evaluación de modelos estadísticos para la elección del modelo que más se ajusta a los datos. |
| Determina el algoritmo que posee mayor ajuste, según las características de la problemática. |
| Identificar los modelos estadísticos existentes de acuerdo al objetivo final del trabajo de minería de datos. | Reconoce los modelos predictivos, según los requerimientos del trabajo de minería de datos. |
| Comprender la aplicación de los modelos en el entorno de programación, de acuerdo, necesidades y características de la organización. | Reconoce funcionalidades que posee Python en la programación y uso del framework de Anaconda. |
| Reconoce los comandos y funciones utilizadas para la aplicación de los modelos en lenguaje Python. |
| Reconocer el uso de los índices de evaluación de algoritmos, de acuerdo a características del modelo estadístico. | Identifica las características de los coeficientes de evaluación, de acuerdo a los resultados posibles del análisis del modelo estadístico. |

1. **Descripción general actividad**

Esta actividad tiene carácter formativo con el objetivo de aplicar lo aprendido hasta el momento relacionado con los modelos predictivos.

Se debe desarrollar en sus grupos.

Disponen de la sesión completa. Al finalizar deberá presentar y explicar los resultados.

**OBJETIVOS**

Seleccionar y aplicar un modelo predictivo a un problema.

**Ejercicio**

Considere el conjunto de datos que se encuentra disponible en comprar\_arrendar.csv se pide seleccionar y aplicar uno de los modelos predictivos revisados.

Su aplicación debe incluir:

* Descripción de dataset
* Revisión de nulos y/o valores atípicos
* Visualización de variables y correlación
* Transformación de variables categóricas y posible escalamiento de variables
* Selección del modelo
* Creación del modelo de Modelos y Comparación de Modelos
* Genere los comentarios en el notebook que permita entender el trabajo realizado

Descripción dataset:

ingresos y gastos son mensuales de 1 personas o 2 si están casados.

trabajo: 0-sin trabajo 1-autonomo 2-asalariado 3-empresario 4-Autonomos 5-Asalariados 6-Autonomo y Asalariado 7-Empresario y Autónomo 8 Empresarios o empresario y autónomo

estado\_civil: 0-soltero 1-casado 2-divorciado

hijos: Cantidad de hijos menores (no trabajan)

comprar: 0-mejor alquilar 1-Comprar casa