|  |  |  |
| --- | --- | --- |
|  | **FORMATO DE GUÍA DE PRÁCTICA DE LABORATORIO / TALLERES / CENTROS DE SIMULACIÓN – PARA DOCENTES** | |
|  | | |
| **CIENCIAS DE LA COMPUTACIÓN** | | INTELIGENCIA ARTIFICIAL  MACHINE LEARNING |
| **TÍTULO PRÁCTICA**: Transformaciones, selección de variables y clasificador KNN | | |
| **OBJETIVO:**  Desarrolla modelos para realizar la clasificación de patrones mediante inteligencia artificial y machine learning  **INSTRUCCIONES**: | | |
| 1. Revisar el contenido teórico del tema | | |
| 1. Profundizar los conocimientos revisando los libros guías, los enlaces contenidos en el material de aprendizaje y la documentación disponible en fuentes académicas en línea | | |
| 1. Desarrollar las actividades de la sección “**ACTIVIDADES POR DESARROLLAR**” | | |
| 1. Subir al AVAC las evidencias de desarrollo de las actividades | | |
| **ACTIVIDADES POR DESARROLLAR**  **Revisar la guía de ejemplo de:**  [**https://genscinet.com/transformacion-clasificacionknn/**](https://genscinet.com/transformacion-clasificacionknn/)  **1. A partir del dataset “Hepatitis”. Describir el dataset.**  Dataset Hepatitis: <https://archive.ics.uci.edu/ml/datasets/Hepatitis>  Describir en la siguiente tabla cada variable, indicar el tipo de variable y técnica o técnicas a utilizar para transformar cada variable. Luego, ejecutar las transformaciones y generar el archivo “DatasetTransformado.csv”.  Basarse en el ejemplo de la siguiente **tabla de diseño de transformaciones**:   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Id** | **Variable** | **Descripción breve** | **Tipo** | **Técnica de transformación a aplicar** | | 1 | sexo | M (masculino) o F (femenino) | Categórica nominal | Sistema binario o coding  Estandarización | | 2 | edad | Años de edad | Numérica discreta | Estandarización | | 3 | nivelformación | Es el nivel de educación:  primaria, secundaria o superior | Categórica ordinal | Ordinal  Estandarización | | … | … |  | … |  |   **2. Modelado sin transformaciones numéricas:** Desarrollar y optimizar un modelo **KNN**. Es decir, entrenar con varias distancias y varios valores de K (cantidad de vecinos) y presentar: a. Una figura que muestre la evolución de Accuracy con respecto a K, b) El mejor valor de K y el Accuracy correspondiente.  **3. Modelado con transformaciones y selección de variables:** A partir del dataset transformado (con transformaciones numéricas), **seleccionar las variables** más relevantes con el propósito de alcanzar una mejor clasificación. Para ello, entrenar con varias distancias y varios valores de K (cantidad de vecinos) y presentar: a. Una figura que muestre la evolución de Accuracy con respecto a K, b) El mejor valor de K y el Accuracy correspondiente.  4. Presentar los resultados en una tabla como la siguiente:   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Modelo con mejores resultados** | **Métrica de distancia** | **Accuracy** | **Precision** | **Recall** | **F1-Score** | | KNN sin transformaciones numéricas |  | … |  |  |  | | KNN con transformaciones numéricas y selección de variables |  | … |  |  |  |   5. Realizar la predicción con un nuevo sample (paciente) con el mejor modelo  6. Presentar conclusiones y referencias en formato APA  7. Presentación ejecutiva de Resultados: comunicar hallazgos iniciales a través de informes o presentaciones que resuman las observaciones y los posibles pasos siguientes en el análisis. Presentar lo siguiente:   * Descripción del dataset, cantidad de variables y observaciones, la variable de estudio Y (objetivo) * Tabla de diseño de transformaciones * Análisis de frecuencia de categorías * Figuras de KNN sin transformaciones * Figuras de KNN con transformaciones y selección de variables * Resultados de calidad de modelos (Accuracy, Precision, Recall, F1-Score, etc.) * Ejemplos de predicción con nuevos datos * Conclusiones finales   **Rúbrica:**  **Una adecuada presentación de resultados (la actividad 7 habilita para la revisión de los cuadernos de Jupyter)**  **Nota: se recomienda utilizar Google Colab**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Criterio | Insatisfactorio (0-2 puntos) | Aceptable (3-5 puntos) | Bueno (6-8 puntos) | Excelente (9-10 puntos) | Puntaje obtenido (el total es el promedio de todos los criterios) | | 1. Descripción de Datos |  |  |  |  |  | | - Variables y Tipo | Se describen algunas variables de manera inexacta o faltan descripciones. | Se describen la mayoría de las variables de manera adecuada, con algunas omisiones menores. | Se describen todas las variables de manera precisa y completa. | Se describen todas las variables de manera precisa y completa, y se incluyen detalles adicionales. |  | | - Técnicas de Transformación | No se proponen técnicas de transformación o las propuestas son incorrectas. | Se proponen técnicas de transformación para algunas variables, pero faltan justificaciones. | Se proponen técnicas adecuadas para la mayoría de las variables con justificaciones razonables. | Se proponen técnicas adecuadas para todas las variables con justificaciones detalladas y razonadas. |  | | 2. Modelado KNN |  |  |  |  |  | | - Optimización de KNN | No se entrena un modelo KNN o se omite la optimización de K. | Se entrena un modelo KNN, pero la presentación de la evolución de Accuracy es insuficiente. | Se entrena un modelo KNN con varios valores de K y se presenta la evolución de Accuracy. | Se entrena un modelo KNN con varios valores de K y se presenta la evolución de Accuracy de manera clara y completa. |  | | - Mejor Valor de K y Accuracy | No se identifica el mejor valor de K o se presenta de manera incorrecta. | Se identifica el mejor valor de K, pero puede haber algunas inconsistencias. | Se identifica el mejor valor de K de manera clara. | Se identifica el mejor valor de K de manera clara y se presenta el Accuracy correspondiente con precisión. |  | | 3. Modelado con Transformaciones y Selección de Variables |  |  |  |  |  | | - Transformaciones | No se realizan transformaciones o son inadecuadas. | Se realizan algunas transformaciones, pero pueden faltar detalles. | Se realizan transformaciones efectivas y adecuadas. | Se realizan transformaciones efectivas y adecuadas con detalles adicionales. |  | | - Selección de Variables | No se lleva a cabo la selección de variables o es insatisfactoria. | Se intenta llevar a cabo la selección de variables, pero puede ser incompleta. | Se lleva a cabo una selección de variables efectiva. | Se lleva a cabo una selección de variables efectiva con justificaciones detalladas. |  | | 4. Presentación de Resultados |  |  |  |  |  | | - Presentación de Resultados | La presentación de resultados es insatisfactoria o está ausente. | La presentación de resultados es insatisfactoria o poco clara. | La presentación de resultados es satisfactoria. | La presentación de resultados es clara y completa. |  | | 5. Predicción con Nuevo Sample |  |  |  |  |  | | - Realización de Predicción | La realización de la predicción es insatisfactoria o está ausente. | Puede haber problemas en la realización de la predicción. | La realización de la predicción es satisfactoria. | La realización de la predicción es eficiente y precisa. |  | | 6. Conclusiones y Referencias |  |  |  |  |  | | - Conclusiones | Las conclusiones son insatisfactorias o están ausentes. | Las conclusiones pueden ser poco claras o insuficientes. | Se presentan conclusiones adecuadas. | Se presentan conclusiones claras y relevantes basadas en los resultados obtenidos. |  | | - Referencias en formato APA | Las referencias en formato APA son insatisfactorias o están ausentes. | Las referencias en formato APA pueden tener algunos errores o faltar detalles. | Las referencias en formato APA son satisfactorias. | Las referencias en formato APA son adecuadas y precisas. |  | | | |
| **RESULTADO(S) OBTENIDO(S)**:  Desarrollan modelos para realizar la clasificación de patrones mediante inteligencia artificial y machine learning  **CONCLUSIONES**:  Los estudiantes desarrollan modelos para realizar la clasificación de patrones mediante inteligencia artificial y machine learning  **RECOMENDACIONES**:   * Revisar la información proporcionada por el docente previo a la práctica. * Haber asistido a las sesiones de clase. * Consultar con el docente las dudas que puedan surgir al momento de realizar la práctica. | | |

***Docente*:** Ing. Remigio Hurtado Ortiz, PhD.

***Firma*: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**