

SILABO

FACULTAD
ESCUELA PROFESIONAL
PROGRAMA DE ESTUDIOS

INGENIERIA ESTADISTICA E INFORMATICA
INGENIERIA ESTADISTICA E INFORMATICA
CARRERA PURA

I. INFORMACIÓN GENERAL

1.1 Identificación Académica

a)	Curso	APRENDIZAJE SUPERVISADO
b)	Código	EST320
c)	Prerequisito	EST314 - APRENDIZAJE NO SUPERVISADO
d)	Número de Horas	03h teóricas, 02h prácticas, Total 05 horas
e)	Créditos	04
f)	Número de Horas virtuales	04
g)	Año y Semestre Académico	2025-II
h)	Ciclo de Estudios	VII
i)	Duración	Del 25 de Agosto al 19 de Diciembre del 2025 (17 semanas)
j)	Área Curricular	Estudios de especialidad
k)	Características del Curso	I + D + I, experiencia pre profesional

1.2 Docente

a)	Apellidos y Nombres	ALEMAN GONZALES LEONID
b)	Condición y Categoría	Docente Nombrado a T.C. Categoría Asociado
c)	Especialidad	Ingeniería Estadística e Informática - Magister en Ingeniería de Software
a)	Apellidos y Nombres	QUISPE YAPO EDGARDO
b)	Condición y Categoría	-
c)	Especialidad	-

1.3 Ambiente donde se realizó el aprendizaje

- a) EI - 102, Laboratorio de Cómputo 02

II. SUMILLA

El curso de Aprendizaje Supervisado forma parte del área curricular de estudios de especialidad y es de naturaleza teórico-práctica. Su propósito es formar competencias en el diseño, implementación y evaluación de modelos estadísticos y computacionales mediante técnicas de aprendizaje automático (machine learning), empleando variables independientes (features) para estimar o clasificar una variable dependiente (output). Se hace especial énfasis en el análisis de datos mediante enfoques de regresión y clasificación, desarrollando tanto fundamentos clásicos como metodologías modernas de alto rendimiento, integrando el uso de lenguajes de programación como Python o R y herramientas de evaluación del desempeño de modelos. El curso se estructura en dos unidades claramente diferenciadas:

Unidad I: Modelos de Aprendizaje Supervisado Basados en Estadística Clásica y Métodos de Interpretación Directa.

Unidad II: Aprendizaje Supervisado de Alto Rendimiento con Redes Neuronales y Técnicas de Ensamblado.

III. PERFIL DEL EGRESADO EN RELACIÓN AL CURSO

CE2. Construye y analiza modelos estadísticos y/o algoritmos computacionales para la toma de mejores decisiones en condiciones de incertidumbre, para el desarrollo de la región y el país, asumiendo una actitud científica y ética, utilizando técnicas estadísticas y/o informáticas referidas a la ciencia de datos.

IV. LOGRO DE APRENDIZAJE DEL CURSO

El estudiante diseña, construye y evalúa modelos de regresión y clasificación, utilizando técnicas de aprendizaje automático con los enfoques clásicos y modernos en problemas reales de competencia de la ingeniería estadística e informática considerando métodos de ajuste, validación, selección de características e interpretabilidad, de manera crítica y ética.

V. TRATAMIENTO DE UNIDADES DIDÁCTICAS

UNIDAD 1		Modelos de Aprendizaje Supervisado Basados en Estadística Clásica y Métodos de Interpretación Directa	
LOGROS DE APRENDIZAJE DE LA UNIDAD Es estudiante, formula, implementa e interpreta modelos clásicos de aprendizaje supervisado, fundamentados en principios estadísticos y matemáticos, para resolver problemas de regresión y clasificación, garantizando la validez del modelo mediante la evaluación rigurosa del ajuste, la generalización y la selección de atributos.			
TIEMPO DE DESARROLLO		Del 25 de Agosto al 27 de Octubre del 2025 (Total 45 horas)	
HORAS DE ENSEÑANZA VIRTUAL/UNIDAD		02	
SEMANAS	CRITERIOS DE DESEMPEÑO		CONOCIMIENTOS

Semana 1	Comprende la metodología y los fundamentos del aprendizaje supervisado, tareas de clasificación y regresión, áreas y definiciones fundamentales.	• Metodología del Aprendizaje Automático, CRISP-DM, KDD • Aprendizaje Automático, estadística, ciencias y áreas relacionadas. • Fundamentos del Aprendizaje Supervisado • Tareas de regresión y clasificación • Definiciones y términos específicos fundamentales.
Semana 2	Analiza y comprende las métricas de evaluación de modelos.	MAE, RMSE, R ² , Accuracy, matriz de confusión, F1-score, validación cruzada, ROC-AUC.
Semana 3	Aplica modelos de regresión lineal y no lineal, a datos reales, analiza coeficientes y errores, y selecciona variantes regularizadas según el ajuste.	• Regresión múltiple lineal • Transformación de variables y espacio de mayor dimensión • Regresión con base radial (RBF), kernel polinómico • Regularización • Regresión Ridge • Regresión Lasso
Semana 4	Entrena modelos logísticos, genera probabilidades y clasifica correctamente con evaluación gráfica y métrica.	• Regresión logística • Log-odds • Función sigmoide • Análisis mediante Curva ROC. RETROALIMENTACIÓN I
Semana 5	Genera un árbol de decisión a partir de datos estructurados y analiza su profundidad, importancia de variables y sobreajuste.	• Árboles de decisión (ID3, C4.5, CART) • Ganancia de información • Poda, visualización.
Semana 6	Implementa algoritmos basados en instancias como clasificador k-NN y evalúa su rendimiento para distintos valores de k.	• Algoritmos basados en instancias • Escalado (normalización, estandarización) • Clasificador k-NN • Criterio de distancias • Selección de k
Semana 7	Aplica LDA a datasets multiclase, explica los supuestos gaussianos del modelo y compara resultados con otros clasificadores.	• Análisis Discriminante Lineal y No Lineal LDA. • Kernel LDA • Comparación con regresión logística
Semana 8	Entrena un modelo SVM y ajusta sus hiperparámetros, comparando el rendimiento según el kernel usado.	• Kernels • Máquinas de Soporte Vectorial (SVM) • SVM lineal y no lineal • Márgenes suaves y parámetros RETROALIMENTACIÓN II
Semana 9	Desarrolla apropiadamente la evaluación y presenta el avance de proyecto de comparación de las técnicas estudiadas a un caso real.	• Evaluación formativa de la unidad • Presentación de avance del trabajo final
PORCENTAJE DE AVANCE ACADÉMICO DE LA UNIDAD: 55%		

* El docente deberá programar “actividades de retroalimentación” los cuales se desarrollarán a lo largo de cada unidad didáctica con el fin de asegurar los adecuados aprendizajes de los estudiantes.

UNIDAD 2	Aprendizaje Supervisado de Alto Rendimiento con Redes Neuronales y Técnicas de Ensamblado
LOGROS DE APRENDIZAJE DE LA UNIDAD El estudiante diseña, implementa, entrena y evalúa modelos supervisados de alto rendimiento utilizando redes neuronales artificiales y técnicas de ensamblado (ensemble), seleccionando arquitecturas, funciones de activación y métodos de regularización adecuados, así como técnicas de validación y ajuste de hiperparámetros, para resolver problemas complejos de clasificación y regresión con datos reales.	
TIEMPO DE DESARROLLO	Del 27 de Octubre al 19 de Diciembre del 2025 (Total 40 horas)

HORAS DE ENSEÑANZA VIRTUAL/UNIDAD		02
SEMANAS	CRITERIOS DE DESEMPEÑO	CONOCIMIENTOS
Semana 10	Construye un MLP simple, ajusta hiperparámetros básicos y evalúa su rendimiento, describe el funcionamiento del gradiente descendente.	- Modelo de redes neuronales artificiales para aprendizaje supervisado - Arquitectura Perceptrón Multicapa - Funciones de activación - Algoritmo de backpropagation - Aprendizaje basado en gradiente descendente
Semana 11	Define, entrena y valida una red profunda, selecciona e implementa el optimizador adecuado comparando su impacto	- Redes densas (DNN), compilación, entrenamiento y evaluación. - Arquitectura de una red profunda - Variantes del gradiente descendente: SGD, Momentum, RMSProp, Adam. - Clasificación de imágenes con Deep Learning básico
Semana 12	Compara árboles individuales y ensambles mediante análisis de varianza y rendimiento.	- Métodos ensemble parte I - Bagging, - Out-Of-Bag error - Random Forest - Bootstrapping - Importancia de variables. - Opcional (Extra Trees)
Semana 13	Ajusta un modelo AdaBoost y evalúa su precisión y evolución iterativa.	- Métodos ensemble parte II - AdaBoost y evaluación iterativa - Clasificadores débiles - Combinación secuencial. - Opcional (Stacking y Voting classifier)
		RETROALIMENTACIÓN III
Semana 14	Implementa y compara distintos algoritmos de boosting con tuning básico	- Gradient Boosting - XGBoost, LightGBM, CatBoost - Algoritmos de gradiente, - Control de overfitting - Eficiencia - Opcional (Grid Search y Optuna)
Semana 15	Evalúa el impacto de la selección, el balanceo, la ingeniería de características en métricas clave del modelo.	- Selección de variables, balanceo de datos e ingeniería de características - Imputación, codificación One-Hot, Label Encoding) - RFE (Recursive Feature Elimination) SelectKBest - Transformaciones no lineales SMOTE - Creación de atributos sintéticos. - Opcional(target encoding, freq encoding, filtros de ruido)
Semana 16	Interpreta decisiones de modelos complejos aplicando SHAP y LIME en un caso práctico.	- Interpretabilidad de modelos complejos - SHAP - LIME - Importancia global/local - Presentación del caso real
		RETROALIMENTACIÓN IV
Semana 17	Integra todas las etapas del proceso supervisado justificando sus decisiones y evaluando el modelo final y superando la evaluación teórica-práctica.	- Evaluación final - Defensa final del proyecto
PORCENTAJE DE AVANCE ACADÉMICO DE LA UNIDAD: 45%		

* El docente deberá programar “actividades de retroalimentación” los cuales se desarrollarán a lo largo de cada unidad didáctica con el fin de asegurar los adecuados aprendizajes de los estudiantes.

VI. ESTRATEGIAS METODOLOGICAS

6.1 De Enseñanza

Actividades sincronicas

- Actividades en video conferencias.
- Implementación de foros de discusión en línea
- Uso de sala de chat en línea.

Actividades asincronicas

- Revisión de foros de información sobre la temática
- Planteamiento y revisión de tareas
- Uso de materiales de lectura
- Proporcionar diapositivas con la temática y videos

6.2 De Aprendizaje

- Análisis de textos y artículos
- Revisión de guías y manuales
- Plataforma en la nube de programación Google COLAB, Jupyter Notebook.
- Debate, análisis y discriminación de ideas.
- Uso de software de lenguajes R y Python y plataformas de programación en Pc y la nube.
- Resolución de casos y exposición de propuestas de solución y análisis de soluciones
- Software IDEs, Pycharm, Jupyter, Spyder, Google Colaboratory
- Resolución de problemas y casos

6.3 De Investigación Formativa

- Revisión de recursos bibliográficos (libros, artículos y tesis) de forma estructurada.
- Preprocesamiento y transformación de datos.
- Presentación de informes y trabajos usando las normas y herramientas de publicación de artículos y reportes de investigación (estructura, partes, gráficos, citas, referencias, similaridad).

6.4 De Responsabilidad Social Universitaria

Analizar problemas de la sociedad a ser abordados, desarrollar la extracción de datos y proponer soluciones buscando aplicar los conocimientos del curso, gestionando un banco de datos y proyectos abiertos.

6.5 De Enseñanza Virtual

Uso de aula virtual, materiales digitales y recursos de aprendizaje de la nube, así como la administración y almacenamiento de recursos virtuales.

VII. MEDIOS Y MATERIALES DIDÁCTICOS

- Uso de plataformas web, máquinas virtuales y servicios web.
- Diapositivas, manuales y guías de aprendizaje digital
- Materiales en video y pizarra en línea.
- Software IDE y framework para Python y R.

VIII. EVALUACIÓN DEL APRENDIZAJE

8.1 Logro de aprendizaje, evidencias de desempeño, ponderación, técnicas e instrumentos de evaluación.

UNIDAD	LOGROS DE APRENDIZAJE	EVIDENCIAS DESEMPEÑO: De acción, objeto o producto (%)	PONDERACIÓN (Obligatorio en base a 100%)	TÉCNICAS	INSTRUMENTOS
1	Es estudiante, formula, implementa e interpreta modelos clásicos de aprendizaje supervisado, fundamentados en principios estadísticos y matemáticos, para resolver problemas de regresión y clasificación, garantizando la validez del modelo mediante la evaluación rigurosa del ajuste, la generalización y la selección de atributos.	Evidencia 1: - Informe - Defensa Evidencia 2: - Examen t/p - Resolución de casos - Implementación	50%	- Revisión de informe - Observación - Examen - Análisis de resultados	- Lista de cotejo - Registro - Rúbrica - Lista de cotejo - Informe
2	El estudiante diseña, implementa, entrena y evalúa modelos supervisados de alto rendimiento utilizando redes neuronales artificiales y técnicas de ensamblado (ensemble), seleccionando arquitecturas, funciones de activación y métodos de regularización adecuados, así como técnicas de validación y ajuste de hiperparámetros, para resolver problemas complejos de clasificación y regresión con datos reales.	Evidencia 3: Informe - Defensa Evidencia 4: Examen t/p - Resolución de casos - Implementación	50%	- Revisión de informe - Observación - Examen - Análisis de resultados	- Lista de cotejo - Registro - Rúbrica - Lista de cotejo - Informe

8.2 Evidencias de aprendizaje del semestre académico.

LOGRO DE APRENDIZAJE DEL CURSO	EVIDENCIAS: ACCIÓN, OBJETO o PRODUCTO	FECHA DE PRESENTACIÓN
--------------------------------	---------------------------------------	-----------------------

El estudiante diseña, construye y evalúa modelos de regresión y clasificación, utilizando técnicas de aprendizaje automático con los enfoques clásicos y modernos en problemas reales de competencia de la ingeniería estadística e informática considerando métodos de ajuste, validación, selección de características e interpretabilidad, de manera crítica y ética.	- Resolución de casos reales comparando diferentes modelos - Presentación del producto aplicación en forma de resultados en artículo o poster.	del 07 al 13 de diciembre 2025
--	---	--------------------------------

• Evidencias que deben ser socializados con la comunidad como parte de la responsabilidad social universitaria. Por la emergencia sanitaria podría optarse por la difusión vía página web y redes sociales.

8.3 Calificación

La fórmula para la obtención del promedio final del curso es la siguiente:

$$\text{Promedio Final} = (50\%)IUPP + (50\%)IIUPP$$

Donde:

IUPP : Primero unidad promedio parcial

IIUPP : Segundo unidad promedio parcial

La fórmula para la obtención del promedio final del curso es la siguiente:

$$\text{Promedio Final} = (50\%)IUPP + (50\%)IIUPP$$

Donde:

IUPP : Primero unidad promedio parcial

IIUPP : Segundo unidad promedio parcial

La fórmula para la obtención del promedio parcial de cada unidad didáctica es la siguiente:

$$\text{Promedio parcial de la unidad} = (30\% (LA1) + 30\% (LA2) + 40\% (LA3)) / (100\%)$$

LA1 = Logro de aprendizaje 1.

LA2 = Logro de aprendizaje 2.

LA3 = Logro de aprendizaje 3.

- La fórmula para la obtención del promedio final del curso es la siguiente:

$$\text{Promedio final} = (I \text{ UPP} + II \text{ UPP}) / 2$$

IX. FUENTES DE INFORMACIÓN

9.1 Bibliográficas

Básica

- Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (3.ª ed.). O'Reilly Media.
- Hastie, T., Tibshirani, R., Friedman, J. (2014). The Elements of Statistical Learning. 2nd. Ed. Springer Series in Statistics.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. Springer.
- Goodfellow I. y Bengio, Y. (2016). Deep Learning, MIT Press. USA.
- Bishop, C. (2006). Pattern Recognition and Machine Learning, Springer. USA.
- Duda, R., Hart, P. y Stork., (2000) Pattern Classification. Wiley Interscience, 2nd. Ed.

Complementarias

- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD
- Chollet, F. (2021). Deep Learning with Python (2.ª ed.). Manning Publications.
- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363.
- Glover, F., Kochenberger G. (1998). Algoritmos e Heurísticas. Ed, Universidad Federal Fluminense. Brasil.
- Stuart Russell & Peter Norvig, (2014), Inteligencia Artificial: Un Enfoque Moderno, 3ra. Ed. Prentice Hall. México.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30.
- Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
- Mitchell, T. M. (1997). Machine Learning. McGraw-Hill.
- Nielsen, M. (2015). Neural Networks and Deep Learning. Determination Press. <https://neuralnetworksanddeeplearning.com>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144).
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). Data Mining: Practical Machine Learning Tools and Techniques (4.ª ed.). Morgan Kaufmann.

Electrónicas

- Algorithm and heuristics (2019) <http://www.cs.sunysb.edu/~algorithm/>
- Python Machine Learning Tutorials <https://realpython.com/tutorials/machine-learning/>

Producción intelectual del docente relacionado con el curso

- Guías de programación con Python para Machine Learning (2019).

Puno, Diciembre del 2025