

# Sumilla

Sitio:	CTIC-UNI Business School Online
Curso:	Machine Learning con Python
Libro:	Sumilla
Imprimido por:	Manuel Castillo-Cara
Día:	lunes, 17 de febrero de 2020, 10:34

# Tabla de contenidos

- 1 Datos generales
- 2 Destinatarios
- 3 Descripción
- 4 Resultados de aprendizaje
- 5 Contenidos
- 6 Estrategias didácticas
- 7 Evaluación
- 8 Docencia
- 9 Bibliografía

# 1 Datos generales

<b>Nombre del curso:</b>	Machine Learning con Python
<b>Año calendario:</b>	2020
<b>Tipo del curso:</b>	Virtual
<b>Prerrequisitos:</b>	Conocimientos intermedios de programación. Conocimientos básicos de estadística inferencial y cálculo de probabilidades.
<b>Unidades:</b>	8
<b>Duración:</b>	8 semanas
<b>Profesor:</b>	Dr. Manuel Castillo-Cara

## 2 Destinatarios

El curso de "Machine Learning con Python" está dirigido a personas que tengan pocos conocimientos de machine learning, conocimientos intermedios del lenguaje de programación Python y que quieran adentrarse a este apasionante mundo de dentro del campo de modelado predictivo.

Además, el curso está diseñado para que cualquier estudiante universitario, investigador o tecnólogo que se encuentre realizando o necesite realizar diferentes experimentos a través de grandes conjuntos de datos para poder sintetizarlos en alguna salida predictiva puedan utilizar los muy diferentes recursos de machine learning que nos pone a nuestra disposición la librería Scikit-Learn y el lenguaje de programación Python.

# 3 Descripción

El curso de “Machine Learning con Python” se centra en un subcampo específico de aprendizaje automático llamado modelado predictivo. Este es el campo del aprendizaje automático que es el más útil en la industria y el cual se utilizar la librería de aprendizaje automático scikit-learn en Python por su gran rendimiento y facilidad en su uso.

A diferencia de las estadísticas, donde los modelos se usan para comprender los datos, el modelado predictivo se enfoca en el desarrollo de modelos que hacen las predicciones más precisas a expensas de explicar por qué se hacen las predicciones.

A diferencia del campo más amplio del aprendizaje automático que podría utilizarse con datos en cualquier formato, el modelado predictivo se centra principalmente en datos tabulares (por ejemplo, tablas de números como en una hoja de cálculo).

## 4 Resultados de aprendizaje

- Aplicar técnicas de análisis y visualización de datos en un conjunto de datos complejo para problemas de machine learning.
- Aplicar técnicas de tratamiento de datos en un conjunto de datos para mejorar la robustez y métrica de salida de los diferentes algoritmos de machine learning.
- Comprender los diferentes mecanismos y técnicas para aplicar analítica predictiva en problemas de machine learning e interpretar la salida obtenida por los modelos de predicción.
- Comprender y analizar la fase del análisis de datos previos al modelado algorítmico en machine learning.
- Realizar modelos algorítmicos robustos con una optimización de sus hiperparámetros para la fase de predicción
- Desarrollar y analizar proyectos de machine learning como regresión, clasificación y multiclase.
- Utilizar librerías específicas de Python como scikit-learn para trabajos de Machine Learning

# 5 Contenidos

## **Módulo 1: Introducción**

1. Conceptos básicos de machine learning.
2. Jupyter Notebook como nuestro entorno de machine learning.
3. Curso rápido de Python.

## **Módulo 2: Análisis de datos**

- 1 Cargar un conjunto de datos.
2. Estadística descriptiva.
3. Visualización de datos.
4. Taller: Trabajo de aplicación de diferentes técnicas analíticas de datos en un conjunto de datos seleccionado por el usuario e interpretar la salida obtenida.
5. Examen tipo test sobre los contenidos del módulo.

## **Módulo 3: Preprocesamiento de datos**

1. Análisis exploratorio de datos.
2. Preprocesamiento de datos.
3. Métodos de remuestreo para estimar la precisión del modelo.
4. Taller: Trabajo de aplicación de diferentes técnicas de análisis y procesamiento de datos de datos en un conjunto de datos seleccionado por el usuario e interpretar la salida obtenida.
5. Examen tipo test sobre los contenidos del módulo.

**Unidad 4: Tratamiento de datos**

1. Evaluación de las métricas.
2. Feature Selection.
3. Feature Importance.
4. Reducción de dimensiones en un dataset.
5. Taller: Aplicación de diferentes técnicas de tratamiento de datos en un conjunto de datos y verificación de su impacto en las métricas algorítmicas.
6. Examen tipo test sobre los contenidos del módulo.

**Unidad 5: Fase de modelado**

1. Algoritmos de Machine Learning.
2. Rendimiento de los algoritmos.
3. Algoritmos Ensamblados
4. Taller: Aplicación de diferentes algoritmos de machine learning en un conjunto de datos e interpretar la salida obtenida, así mismo, verificar el algoritmo que tenga mejor comportamiento.
5. Examen tipo test sobre los contenidos del módulo.

**Unidad 6: Fase de optimización y forecasting**

1. Pipelines
2. Procesamiento de datos avanzado
3. Configuración de hiperparámetros
4. Guardado e integración del modelo
5. Taller: Una vez seleccionados los algoritmos candidatos a modelo realizar una optimización de estos a través de la configuración de sus hiperparámetros.
6. Examen tipo test sobre los contenidos del módulo.

**Unidad 7: Proyectos de Machine Learning**

1. Trabajar un proyecto de clasificación multiclase
2. Trabajar un proyecto de regresión.
3. Trabajar un proyecto de clasificación binaria.
4. Proyecto: Realizar un proyecto completo analizando todas las fases estudiadas en los diferentes módulos.



# 6 Estrategias didácticas

Curso Virtual el curso es virtual para poder llevarlo a cabo se tiene el siguiente esquema:

1. Lectura del material el Aula Virtual. Los participantes debe leer y revisar los contenidos teóricos que se tienen en el Aula Virtual, correspondiente a cada unidad.
2. Visualización de Vídeos. Como ayuda al aprendizaje se tienen vídeos para cada unidad, los cuales deben ser visualizados por el alumnado.
3. Autoaprendizaje. Se debe resolver el material para refuerzo y aplicación de los contenidos teóricos/prácticos antes de la evaluación.

## **Software de trabajo**

- Se utilizara el software Jupyter Notebook como nuestro entorno de programación.
- Se utilizará el software libre Scikit-Learn.

## **Foro de consultas**

- Este espacio está destinado para que los estudiantes formulen sus preguntas con respecto a la temática desarrollada y el docente tutor será el responsable de absolver sus interrogantes.

## **Actividad virtual**

- Sesiones de videoconferencias
- Análisis de casos
- Foros de discusión
- Trabajos parciales de los módulos en foros
- Examen tipo test.
- Proyecto final del curso
- Lecturas comentadas y

- Búsquedas de información científica.

# 7 Evaluación

La evaluación del aprendizaje será un proceso continuo donde los estudiantes tienen la oportunidad de ir trabajando con el software de machine learning practicando lo expuesto en la parte de teoría.

Se tomará en cuenta una evaluación desde el módulo 3 al módulo 6 y en la que cada unidad tendrá dos notas:

- **Teoría:** Tiene un ponderación del 30% de la nota final (el total de todas la unidades). Consistirá en examen tipo test online de los conceptos teóricos.
- **Práctica:** Tiene una ponderación del 20% de la nota final (el total de todas la unidades). Se pondrán a prueba los conceptos teóricos explicados en los videotutoriales de manera que el alumno trabajará un proyecto de cada unidad.
- **Proyecto:** Tiene una ponderación del 50% de la nota final. Se realizará un proyecto final del curso el cual deberá poner a prueba todos los contenidos vistos en los diferentes módulos.

Los exámenes se realizarán en 5 de las 7 unidades siendo la la nota final la nota media de estas.

Asimismo, los alumnos deberán trabajar un proyecto consolidado con unos datos propios con todas las fases que lleva un problema de modelado predictivo con Python.

## ESCALA DE NOTAS

Calificativos y equivalencias en términos literales

Nota	Nota Literal
17 – 20	A
15– 16	B

13– 14	C
11– 12 ≤ de 10	D E (desaprobado)

- La nota final aprobatoria del curso será de once (11.00). Se expresará en escala vigesimal con dos decimales, no habrá redondeo al valor inmediato superior o inferior.
- La nota de los estudiantes que no rindan la evaluación dentro del plazo establecido para cada unidad será cero (0).

# 8 Docencia



## Dr. Manuel Castillo-Cara

Profesor Principal a Tiempo Completo en la Escuela de Ciencia de la Computación de la Facultad de Ciencias y jefe del laboratorio de investigación Tecnologías Ubicuas Inteligentes - Smart Cities (UITL-SCi) en CTIC - UNI, PhD en Tecnologías Informática Avanzadas por la Universidad de Castilla-La Mancha (UCLM) de España y calificado como investigador CONCYTEC en Perú en la categoría Carlos Monge III en la base de datos CTIVitae.

Gran conocimiento sobre Agricultura de precisión y Smart Cities, en las áreas de toma de decisiones inteligentes a través de una plataforma Internet of Things - Fog Computing y con diferentes fuentes de financiamientos internacional para la ejecución de proyectos de investigación.

Experiencia en redes de sensores y transmisión de señales, aplicaciones de teléfonos móviles y análisis/tratamiento de datos generados tanto por sensores como por dispositivos inalámbricos, haciendo predicciones con diferentes técnicas algorítmicas de Machine learning, así como el alertas

en tiempo real con procesamiento de eventos complejos.

Entre los trabajos asociados a mi laboratorio de investigación se encuentran proyectos para la prevención de violencia de género y seguridad ciudadana y vial, Desarrollo de infraestructuras y sistemas para las Smart Cities y proyectos de índole social Popularización por la ciencia en niños y adolescentes a través de clases de robótica educativa en zonas vulnerables de Perú.

## 9 Bibliografía

- Pattern Recognition and Machine Learning. Christopher M. Bishop. Ed. Springer.
- Machine Learning. A probabilistic perspective. Kevin P. Murphy. Ed. The MIT Press.
- Principles of Data Mining. D. Hand & P. Smyth & H. Mannila. ED. The MIT press.
- Introducción a la minería de datos. J. Hernández Orallo, M.J. Ramírez Quintana, C. Ferri Ramírez. Ed. Pearson.
- Data Mining: Concepts and techniques. Jiawei Han y Micheline Kamber. Ed. Morgan Kauffmann.