

# Programa “Capacitaciones en Inteligencia Artificial 2024”

En el marco de la convocatoria a Capacitaciones IA 2024 brindado por el Centro de Posgrados y Actualización Profesional en Informática (CEPAP) de la Facultad de Ingeniería de Udelar y organizado por Agesic, se estarán dictando diversos cursos, para lo cual invitamos a los funcionarios del Estado a inscribirse.

Estos cursos son: "Introducción a la estadística usando software", "Análisis de textos", "Análisis en series temporales", "Modelado y procesamiento de grandes volúmenes de datos" y "Fundamentos y herramientas para minería de procesos de negocio".

El objetivo de la participación en estos cursos es que puedan aplicar en los organismos el conocimiento adquirido y poder acompañarlos en la aplicación de lo aprendido en sus proyectos.

Estos cursos son presenciales en Facultad de Ingeniería, Udelar y cuyas fechas y horario de dictado se detallan en la sección de cada curso más adelante en este documento.

## Postulación y selección

Debido a que los cupos son limitados y cada curso requiere de conocimientos en algunas áreas específicas, se realizará una selección entre todos los interesados que se hayan postulado, según cuenten con dichos requisitos y/o experiencia.

Para la postulación será necesario, aclarar la motivación personal, además de sus conocimientos y experiencia, que será solicitado en el formulario correspondiente.

Luego, para aquellos que queden seleccionados, se solicitará una carta o email con el aval de su jefe directo en el organismo en el que se desempeña. La postulación a cada curso es individual, para ello se presentan los links correspondientes en el cuadro de la última página de este documento.

## Público Objetivo

El curso está destinado integrantes de equipos de organismos del Estado, profesionales de informática e interesados en las áreas afines de cada curso. El perfil puede ser técnico informático o de gestión y que sean integrantes de equipos de organismos públicos del Estado que realicen actividades vinculadas a las temáticas correspondientes.

**Requisitos deseables:** conocimientos y experiencia en procesos de negocio, modelado, ejecución y sistemas de información basados en procesos, conocimientos de programación, manejo básico de R y conocimientos de probabilidad y estadística. En la información de cada curso se detalla los requisitos deseables específicos.

## Objetivo general

Introducir los aspectos básicos de cada curso. Aplicar los conceptos y técnicas aprendidas en la resolución de los ejercicios que se van proponiendo durante el curso y en la prueba final de cada uno.

## Contenidos del programa

A continuación, se describe el contenido de cada uno de los cursos:

### Curso - Introducción a la estadística usando software

Objetivo: introducir los aspectos básicos de la inferencia estadística clásica. Aplicar dichas técnicas a conjuntos de datos reales e interpretar los resultados obtenidos. Acercar el estudiante al empleo de los paquetes para el análisis estadístico (MATLAB, OCTAVE).

Temáticas abordadas:

- Repaso de Estadística Descriptiva. Histograma. Indicadores de media y de dispersión. Relaciones entre variables cualitativas, variables cualitativas y cuantitativas, variables cuantitativas.
- Repaso de Probabilidad. Variable aleatoria. Independencia. Momentos. Distribuciones usuales. Simulación de variables.
- Estimación puntual. Métodos de estimación. Test de hipótesis. Nociones generales. Test sobre la media en una población normal.
- Estadística no paramétrica. Pruebas de aleatoriedad, ajuste y comparación.
- Relaciones entre variables. Modelos lineales.

Conocimientos exigidos: conocimientos de probabilidad y estadística.

Conocimientos deseables: conocimientos de programación.

Inicio: 15-04-2024

Fin: 24-05-2024

Horarios: martes y jueves de 18 a 21 h.

### Curso - Análisis de textos

Objetivo: el objetivo del curso es introducir los conceptos, los métodos y las aplicaciones principales del área Procesamiento de Lenguaje Natural, haciendo énfasis en problemas y enfoques actuales, basados principalmente en técnicas de Aprendizaje Automático. Se presentará un panorama completo de los desafíos que involucra el área y se experimentará con las herramientas y los recursos habituales, incluyendo alguna práctica orientada a problemas de dimensiones reales de gran desarrollo en la actualidad, como Análisis de Sentimientos y Extracción de Información en Redes Sociales.

Temáticas abordadas:

- Introducción al Procesamiento de Lenguaje Natural Conceptos generales, niveles clásicos de análisis.
- Aplicaciones más comunes del área: traducción automática, resumen automático, extracción de información, análisis de textos de prensa y de redes sociales.
- Herramientas y recursos usuales: analizadores lingüísticos, corpus.
- Conceptos generales de gramática del español: definiciones básicas: oración y sintagma, categorías gramaticales (léxicas y sintagmáticas), funciones sintácticas y semánticas. Análisis gramatical básico de oraciones.
- Tratamiento previo y procesamiento usual de los textos: obtención de textos. Pre-procesamiento: limpieza, tokenización, lematización, stemming, pos-tagging.
- Análisis sintáctico y semántico: recursos disponibles, aplicaciones y problemas. Recursos léxicos. Corpus anotados.
- Métodos clásicos y actuales para el PLN: métodos simbólicos (reglas). Métodos de aprendizaje automático basados en atributos.
- Modelos de datos de desarrollo reciente: representaciones distribuidas de palabras y aprendizaje profundo.
- Problemas actuales del área: análisis de sentimientos. Extracción de información.

Conocimientos exigidos: programación.

Conocimientos deseables: lenguajes formales, aprendizaje automático, probabilidad y estadística, álgebra.

Inicio: 27-05-2024

Fin: 21-06-2024

Horarios: martes y jueves de 18 a 21 h.

### **Curso - Análisis en series temporales**

Objetivo: introducir los aspectos básicos de la modelación de series de tiempo, así como de la predicción de valores futuros.

Temáticas abordadas:

- Conceptos fundamentales. Análisis descriptivo, componentes de tendencia, estacionalidad y componente irregular o aleatoria (3h)
- Procesos estocásticos estacionarios y no estacionarios, función de autocorrelación (3h)
- Modelos autorregresivos (AR), modelos de medias móviles (MA), mixtos (ARMA). Ajuste del modelo (estimación de parámetros). Selección y validación de los modelos. Predicción puntual y por intervalos de valores futuros. (8h)
- Series de memoria corta y de memoria larga. Modelos ARIMA y ARFIMA (3h).
- Redes neuronales. Aplicación a la predicción de series de tiempo (5h).

Conocimientos exigidos: introducción a la probabilidad y estadística.

Conocimientos deseables: manejo básico de R (no excluyente).

Inicio: 24-06-2024

Fin: 19-07-2024

Horarios: lunes, miércoles y viernes 18 a 21 h.

### **Curso - Modelado y procesamiento de grandes volúmenes de datos**

Objetivo: el objetivo de esta asignatura es brindar una visión general sobre estos nuevos modelos de gestión de datos, analizando las fortalezas y debilidades de cada uno, y buscando así enriquecer la mirada y el espíritu crítico sobre estas estrategias que llegaron para quedarse. Durante el curso se hará énfasis en los modelos de datos subyacentes y los lenguajes de consulta asociados. Al final del curso el estudiante contará con elementos para decidir cuál es el modelo y la estrategia de gestión de datos que mejor se ajusta a su problema.

Temáticas abordadas:

- Introducción
  - Perspectiva histórica y evolución de los sistemas de bases de datos.
  - Modelo conceptual
- El modelo relacional
  - Introducción al modelo relacional
  - Restricciones de integridad
  - Propiedades ACID
  - SQL
- Bases de datos de documentos
  - Introducción al modelado basados en documentos
  - Aspectos técnicos de MongoDB
  - Consultas en MongoDB
- Bases de datos de grafos
  - ¿Qué es un grafo?
  - Introducción a los modelos de grafos
  - Aspectos técnicos de Neo4j
  - Neo4j Cypher
- BigData y el ecosistema Apache Hadoop
  - Definición de BigData
  - Principales componentes del ecosistema Apache Hadoop (Spark, Cassandra, Hive, Pig)
  - Aspectos técnicos de Hadoop
  - Hadoop MapReduce

Conocimientos exigidos: ninguno.

Conocimientos deseables: programación.

Inicio: 09-09-2024

Fin: 01-11-2024

Horarios: martes y jueves de 18 a 21 h.

## Curso - Fundamentos y herramientas para minería de procesos de negocio

Objetivo: brindar una visión general de los temas asociados a la minería de procesos de negocio como elemento clave en la gestión de procesos de negocio y su mejora. Introducir el ciclo de vida de los procesos de negocio, desde su modelado, implementación, ejecución y evaluación, de forma de comprender el papel fundamental que juega el modelado como aspecto clave para gestionar una organización dirigida por procesos, y la minería de procesos como parte integral de las fases de modelado y evaluación. Comprender los conceptos básicos y enfoques para la minería de procesos de negocio: descubrimiento de modelos, conformidad de modelos y extensión de modelos. Brindar una visión general del descubrimiento de modelos de procesos de negocio desde los datos de ejecución real de los procesos (logs de eventos) así como algoritmos existentes y sus resultados. Introducir los principales desafíos para la obtención de logs de eventos, incluyendo calidad de datos, que permitan la minería de procesos. Identificar casos de uso asociados a la ejecución de procesos, medidas de ejecución, soporte organizacional, entre otros. Adquirir habilidades prácticas para el análisis de la ejecución de procesos de negocio con minería de procesos en herramientas open source como ProM y otras como Disco, etc.

Temáticas abordadas:

- Introducción a la minería de procesos
  - Gestión de procesos de negocio (conceptos, ciclo de vida)
  - Minería de procesos de negocio (descubrimiento, conformidad, extensión, log de eventos)
  - Análisis de datos organizacionales vs análisis de ejecución de procesos
  - Introducción a la minería de procesos con ProM y Disco (laboratorio)
- Descubrimiento de procesos
  - Desafíos para el descubrimiento de procesos
  - Métricas para el descubrimiento de procesos (4 fuerzas)
  - Algoritmo Alpha para descubrimiento de procesos
  - Evaluación de algoritmos de descubrimiento (benchmark)
  - Descubrimiento de procesos con ProM (laboratorio)
- Comprobación de conformidad y extensión de procesos
  - Desvíos típicos en la ejecución de procesos
  - Métricas para la comprobación de conformidad
  - Técnicas para comprobación de conformidad (footprints, token, replay, alignments)
  - Perspectivas de extensión (organizacional, tiempo)
  - Comprobación de conformidad y extensión de procesos con ProM (laboratorio)
- Obtención de datos para minería de procesos
  - Fuente y selección de datos de ejecución de procesos
  - Principales desafíos en la extracción de datos
  - Elementos de calidad de datos del log de eventos
  - Log de eventos en Disco y ProM (laboratorio)
- Aspectos metodológicos y soporte operacional
  - Metodología para proyectos de minería de procesos

- Medidas de ejecución asociadas a los modelos y logs de ejecución
- Análisis de ejecución de procesos (dashboards, minería de procesos)
- Soporte operacional (detección, predicción, recomendación)
- Aplicación a procesos de tipo “lasagna” y “spaguetti”
- Soporte operacional en ProM (laboratorio)

Conocimientos exigidos: ninguno.

Conocimientos deseables: conocimientos sobre procesos de negocio, modelado, ejecución y sistemas de Información basados en procesos.

Inicio: 11-11-2024

Fin: 06-12-2024

Horarios: lunes, miércoles y viernes de 18 a 21 h.

### **Desarrollo del programa**

El programa lo dictarán docentes del Centro de Posgrados de la Facultad de Ingeniería (CPAP) de la Universidad de la República, todos de manera presencial en la institución.

### **Dedicación y evaluación**

Se recomienda reservar horas adicionales de estudio y práctica fuera del horario de clase para garantizar un aprendizaje óptimo y alcanzar los objetivos del curso con éxito.

#### **Curso - Introducción a la estadística usando software**

- Clases presencial: 36 horas (12 clases)
- Evaluación:
  - Prueba final escrita, individual y presencial.
  - Entrega de un trabajo práctico
- Dedicación fuera de clase: SI

#### **Curso - Análisis de textos**

- Clases presencial: 24 horas (8 clases)
- Evaluación:
  - Prueba final escrita, individual y presencial
  - Informe sobre proyecto final grupal
- Dedicación fuera de clase: SI

#### **Curso - Análisis en series temporales**

- Clases presencial: 36 horas (12 clases)
- Evaluación:
  - Prueba final escrita, individual y presencial.
  - Entrega de trabajo práctico.

- Dedicación fuera de clase: Sí

#### Curso - Modelado y procesamiento de grandes volúmenes de datos

- Clases presencial: 48 horas (16 clases)
- Evaluación:
  - Prueba final presencial individual.
  - Entregas de ejercicios prácticos en grupos
- Dedicación fuera de clase: SI

#### Curso - Fundamentos y herramientas para minería de procesos de negocio

- Clases presencial: 36 horas (12 clases)
- Evaluación:
  - Prueba final presencial individual.
  - Participación en clase y trabajo final grupal
- Dedicación fuera de clase: SI

#### Materiales del curso

Se pondrán a disposición de los participantes los materiales del curso y los sugeridos por los docentes (diapositivas del curso y ejercicios).

#### Calendario detallado

Curso	Inicio	Fin	Horarios
Introducción a la estadística usando software	15-04-2024	24-05-2024	Martes y jueves de 18 a 21 h.
Análisis de textos	27-05-2024	21-06-2024	Martes y jueves de 18 a 21 h.
Análisis en series temporales	24-06-2024	19-07-2024	Lunes, miércoles y viernes de 18 a 21 h.
Modelado y procesamiento de grandes volúmenes de datos	09-09-2024	01-11-2024	Martes y jueves de 18 a 21 h.
Fundamentos y herramientas para minería de procesos de negocio	11-11-2024	06-12-2024	Lunes, miércoles y viernes de 18 a 21 h.