

TEMARIO

TECNICO EN MINERIA DE DATOS

Informática

Computadoras y sistemas informáticos. Representación y procesamiento de la información. Computadoras. Hardware y software. CPU. Memorias. Dispositivos de almacenamiento. Sistemas operativos. Periféricos y utilitarios. Redes. Introducción a la programación. Algoritmos, lenguajes y programación estructurada.

Arquitectura de Computadoras

Características generales de una computadora, Sistemas numéricos. Representación de los datos a nivel máquina. Álgebra de Boole. Arquitectura y organización de computadoras. Esquema de funcionamiento. El lenguaje ensamblador: programación Assembler, pilas e interrupciones. Buses y unidad de control. Interfaces y periféricos. Memorias y Microprocesadores. Jerarquía de memoria. Organización funcional. Arquitecturas no Von Neumann. Arquitecturas multiprocesadores.

Introducción a las Comunicaciones

Arquitectura de redes de datos, estándares, Modelos de referencia. Modelos básicos de transmisión. Tecnologías de transmisión. Señales, ruido, distorsión y errores. Canales y capacidad. Modulación de datos. Codificación. Multiplexación. Tecnologías de redes LAN. Ethernet, paquetes, direccionamiento, control de acceso al medio, bridging, vlans y trunking. Wifi: paquetes, control de acceso al medio, access points, seguridad, alternativas para la capa física. Tecnologías para redes WAN: tecnologías actuales para redes de acceso, tecnologías actuales para backbone. TCP/IP. Capa de red: datagramas, direccionamiento, routers, tablas de ruteo, algoritmos de ruteo. Capa de transporte: datagramas UDP, conexiones TCP, segmentos, multiplexación de aplicaciones, control de congestión, control de flujo, resolución de pérdidas y reordenamiento. Mensajes de control. Aplicaciones y servicios básicos de comunicaciones: DNS, Correo electrónico, HTTP, FTP. Analizadores de protocolos.

Sistemas Operativos

Conceptos de los mecanismos de Sistemas Operativos, sus procesos, memoria y manejo de archivos. Introducción a los sistemas operativos. Administración de procesos. Planificación de procesos. Administración de recursos compartidos, sincronización, comunicación entre procesos. Abrazo mortal (deadlock, inter bloqueo o bloqueo mutuo). Administración de memoria. Memoria virtual. Administración de dispositivos de entrada-salida. Administración de archivos. Protección. Introducción a los sistemas distribuidos.

Bases de datos.

Introducción a las bases de datos. Características del software orientado a bases de datos. Modelos de datos, esquemas e instancias. Arquitectura de una base de datos. Conceptos de modelado. Modelo Entidad/Relación. Notación y diagramas. Modelado de problemas reales. Modelo relacional. Conceptos y restricciones. Definición y actualización de relaciones. Claves primarias y claves foráneas. Diseño de bases de datos. Dependencias funcionales. Proceso de normalización hasta forma normal de Boyce-Codd. Lenguajes de consulta. Algebra relacional. Lenguaje SQL. Consultas y actualización de datos en lenguaje SQL. Conexión con una base de datos. Usuarios y permisos de acceso.

Modelos de base de datos: Modelo relacional. Diseño de base de datos: Normalización. Índices. Consultas SQL. Administración de base de datos: Seguridad. Backups. Recuperación de desastres. Aplicaciones de bases de datos: Sistemas de información. E-commerce. Big data.

Probabilidad

Introducción a la Probabilidad: Definición y objetivos. Experimento aleatorio, espacio muestral y evento. Conceptos Básicos: Probabilidad de un evento: definición y propiedades. Eventos complementarios, mutuamente excluyentes e independientes. Probabilidad Condicional: Definición y fórmula. Reglas de Probabilidad: Ley de la suma (para eventos mutuamente excluyentes y no mutuamente excluyentes). Ley del producto. Técnicas de Conteo: Principio fundamental de conteo. Permutaciones y combinaciones. Distribuciones de Probabilidad Discretas: Concepto y características. Distribución uniforme discreta. Distribución binomial. Distribución geométrica. Distribución de Poisson.

(Otros modelos pueden ser incluidos dependiendo del nivel y enfoque del curso).
Distribuciones de Probabilidad Continuas: Concepto y características. Distribución uniforme continua. Distribución normal (o gaussiana). Propiedades y aplicaciones de la distribución normal. Variables Aleatorias: Definición y tipos (discretas y continuas). Esperanza, varianza y desviación estándar de una variable aleatoria.

Minería de datos

Introducción a la minería de datos: Conceptos básicos de minería de datos. Tipos de minería de datos. Algoritmos de minería de datos. Exploración de datos: Limpieza de datos. Visualización de datos. Reprocesamiento de datos. Minería de patrones: Agrupamiento. Asociación. Clasificación. Regresión. Aplicaciones de minería de datos: Marketing. Finanzas. Salud. Ciencia.

Visualización de datos

Conceptos básicos de visualización de datos: Definición de visualización de datos. Tipos de visualización de datos. Principios de visualización de datos. Herramientas de visualización de datos: Software de visualización de datos. Librerías de visualización de datos. Aplicaciones de visualización de datos: Comunicación de datos. Comprensión de datos.

Conceptos Básicos de Estadística

Definición de Estadística: diferencia entre estadística descriptiva y estadística inferencial. Tipos de Datos. Datos cuantitativos (continuos y discretos) y cualitativos (nominales y ordinales). Niveles de medición: nominal, ordinal, intervalo y razón. Fuentes de Datos: Datos primarios y secundarios. Observacionales y experimentales. Muestreo: Población y muestra. Métodos de muestreo: aleatorio simple, sistemático, estratificado, por conglomerados, entre otros. Frecuencias y Distribuciones: Tablas de frecuencia. Histogramas, polígonos de frecuencia y ojivas. Frecuencia acumulada y relativa. Medidas de Tendencia Central: Media (promedio). Mediana. Moda. Medidas de Dispersión: Rango. Varianza y desviación estándar. Coeficiente de variación. Posición: Cuartiles, deciles, percentiles. Relación y Dependencia: Correlación y causalidad. Coeficiente de correlación.

Gráficos Estadísticos Básicos: Diagramas de barras, circulares (pastel o tarta) y de dispersión. Boxplots (diagramas de caja y bigotes). Probabilidad Básica: Experimentos, eventos y espacio muestral. Probabilidad condicional y eventos independientes. Teorema de Bayes (introducción). Distribuciones Básicas: Distribución uniforme, binomial y normal (introducción).

Estadística descriptiva, introducción a la estadística descriptiva: definición y objetivos. Diferencia entre estadística descriptiva y estadística inferencial. Tipos de Datos: Datos cuantitativos (continuos y discretos) y cualitativos (nominales y ordinales).

Algoritmos y estructuras de datos

Conceptos básicos de algoritmos: Definición de algoritmo. Notación asintótica. Complejidad de algoritmos. Diseño de algoritmos. Estructuras de datos fundamentales: Listas enlazadas. Pilas. Colas. Árboles. Grafos. Algoritmos fundamentales: Búsqueda. Ordenamiento. Recursividad. Programación dinámica. Aplicaciones de algoritmos y estructuras de datos: Inteligencia artificial. Computación gráfica. Criptografía. Ciencias de la computación.