Logotipo

Descripción generada automáticamente con confianza media

INSTITUTO TECNOLÓGICO DE CULIACÁN

ING. EN SISTEMAS COMPUTACIONALES

TAREA 3

ALUMNO:

CHRISTIAN NATANAEL AYON ZAZUETA

MATERIA:

INTELIGENCIA ARTIFICIAL

MAESTRO:

ZURIEL DATHAN MORA FELIX

HORARIO

9:00-10:00

FECHA

16/02/2025

**Cómo Construir un Sistema de Recomendación, Tecnologías y Frameworks de Desarrollo**

**Conceptos Básicos:**  
Un sistema de recomendación es una herramienta que sugiere ítems (productos, películas, música, etc.) a los usuarios basándose en sus preferencias, comportamientos y datos históricos. Los sistemas de recomendación se clasifican en tres tipos principales:

* **Filtrado Colaborativo:** Basado en la similitud entre usuarios o ítems.
* **Filtrado Basado en Contenido:** Basado en las características de los ítems y las preferencias del usuario.
* **Sistemas Híbridos:** Combinación de filtrado colaborativo y basado en contenido.

**Tecnologías y Frameworks de Desarrollo:**

* **Python:** Es el lenguaje más popular para desarrollar sistemas de recomendación debido a su amplia gama de librerías.
  + **Librerías:**
    - **Scikit-learn:** Para implementar algoritmos básicos de machine learning.
    - **Surprise:** Especializada en sistemas de recomendación.
    - **TensorFlow/PyTorch:** Para modelos más complejos y deep learning.
    - **LightFM:** Para modelos híbridos.
    - **Pandas/Numpy:** Para manipulación de datos.
* **Apache Spark:** Para manejar grandes volúmenes de datos y realizar recomendaciones en tiempo real.
  + **MLlib:** Librería de machine learning de Spark que incluye algoritmos de recomendación.
* **Frameworks Específicos:**
  + **Amazon Personalize:** Servicio gestionado de AWS para construir sistemas de recomendación.
  + **Google Recommendations AI:** Servicio gestionado de GCP para recomendaciones personalizadas.

**Proceso de Construcción:**

1. **Recopilación de Datos:**
   * **Datos de Usuarios:** Información demográfica, preferencias, historial de compras, etc.
   * **Datos de Ítems:** Características de los productos, categorías, precios, etc.
   * **Interacciones:** Ratings, clicks, compras, etc.
2. **Preprocesamiento de Datos:**
   * **Limpieza de Datos:** Eliminar datos inconsistentes o faltantes.
   * **Normalización:** Escalar datos para que tengan un rango común.
   * **Transformación:** Convertir datos categóricos en numéricos.
3. **Selección de Algoritmo:**
   * **Filtrado Colaborativo:** Basado en la similitud entre usuarios o ítems.
   * **Filtrado Basado en Contenido:** Basado en las características de los ítems.
   * **Modelos Híbridos:** Combinación de ambos enfoques.
4. **Entrenamiento del Modelo:**
   * **División de Datos:** Dividir los datos en conjuntos de entrenamiento y prueba.
   * **Entrenamiento:** Ajustar el modelo a los datos de entrenamiento.
   * **Validación:** Evaluar el modelo con datos de prueba.
5. **Evaluación y Optimización:**
   * **Métricas de Evaluación:** Precisión, recall, F1-score, RMSE, etc.
   * **Optimización:** Ajustar hiperparámetros y técnicas de regularización.
6. **Despliegue:**
   * **Integración:** Integrar el modelo en la aplicación o plataforma.
   * **Monitoreo:** Monitorear el rendimiento del modelo en producción.

**Herramientas en AWS/GCP para Desarrollar Sistemas de Recomendación**

**AWS (Amazon Web Services):**

* **Amazon Personalize:** Servicio gestionado que utiliza machine learning para crear recomendaciones personalizadas. Permite entrenar modelos con datos históricos y generar recomendaciones en tiempo real.
  + **Características:** Integración con S3, DynamoDB, y otros servicios de AWS.
  + **Algoritmos:** Incluye algoritmos como User-Personalization, Related Items, y Personalized Ranking.
* **SageMaker:** Servicio de machine learning que permite construir, entrenar, y desplegar modelos personalizados.
  + **Algoritmos:** Incluye algoritmos pre-entrenados para recomendación como Factorization Machines.
* **DynamoDB:** Base de datos NoSQL que puede ser utilizada para almacenar y recuperar datos de usuarios e ítems en tiempo real.

**GCP (Google Cloud Platform):**

* **Recommendations AI:** Servicio gestionado que utiliza modelos de machine learning para generar recomendaciones personalizadas.
  + **Características:** Integración con BigQuery, Cloud Storage, y otros servicios de GCP.
  + **Algoritmos:** Utiliza modelos avanzados de deep learning para recomendaciones.
* **AI Platform:** Servicio de machine learning que permite entrenar y desplegar modelos personalizados.
  + **Algoritmos:** Soporta modelos personalizados y pre-entrenados.
* **BigQuery:** Servicio de almacenamiento y análisis de datos que puede ser utilizado para procesar grandes volúmenes de datos y generar recomendaciones.

**Algoritmos para Optimizar Recursos**

**Algoritmos de Recomendación:**

* **Matrix Factorization:** Descompone la matriz de interacciones usuario-ítem en factores latentes para predecir preferencias.
  + **Ejemplo:** Alternating Least Squares (ALS).
* **K-Nearest Neighbors (KNN):** Encuentra los ítems o usuarios más similares basándose en métricas de distancia.
  + **Ejemplo:** User-based o Item-based KNN.
* **Deep Learning:** Utiliza redes neuronales para capturar patrones complejos en los datos.
  + **Ejemplo:** Neural Collaborative Filtering (NCF).
* **Content-Based Filtering:** Utiliza características de los ítems para recomendar ítems similares a los que el usuario ha interactuado previamente.

**Optimización de Recursos:**

* **Algoritmos de Optimización:**
  + **Gradient Descent:** Utilizado para optimizar los parámetros de los modelos de recomendación.
  + **Stochastic Gradient Descent (SGD):** Variante más eficiente para grandes conjuntos de datos.
* **Técnicas de Muestreo:**
  + **Negative Sampling:** Reduce la complejidad computacional al muestrear solo un subconjunto de ítems no interactuados.
* **Distributed Computing:**
  + **Apache Spark:** Permite distribuir el procesamiento de datos y entrenamiento de modelos en clusters.