**Ad Click Prediction Project**

This project is designed to explore the relationship between user behavior and online advertising performance. It focuses on predicting whether a user will click on an advertisement, using advanced machine learning models. The dataset used in this project contains key insights, including user demographic details (such as age, gender, and income), browsing history, interaction patterns, and specifics about the displayed advertisements (like placement, size, and content type).

By analyzing these factors, the project aims to create highly accurate binary classification models to forecast user interactions with ads. Such predictive models can provide valuable insights for online marketers, advertisers, and businesses by:

• Optimizing ad placement and targeting.

• Reducing advertisement spending on uninterested audiences.

• Improving the return on investment (ROI) in digital marketing campaigns.

• Enhancing personalization for users by serving them more relevant content.

The project explores several machine learning techniques, such as logistic regression, decision trees, random forests, and neural networks, to identify the most efficient approach for predicting ad clicks. In addition to model development, the project emphasizes feature engineering and data preprocessing to handle missing values, outliers, and imbalanced data effectively.

This study has broad applications for digital marketers, e-commerce platforms, social media networks, and any organization looking to boost their online advertising strategy.

The Ad Click Prediction task typically uses machine learning techniques to analyze user behavior and predict if a user will click on an ad. Here are a few key studies and approaches related to this area:

1. Feature Engineering and User Profiling: Studies often focus on designing features based on user behavior, device type, geographical location, and timing. For example, creating features based on session frequency, recency, and interaction history is common.

2. Use of Machine Learning Models: Widely used models for ad click prediction include Logistic Regression, Random Forests, Gradient Boosting Machines, and newer neural networks like deep feedforward networks. The choice of model often depends on the data’s complexity and the computational resources available.

3. Deep Learning: Advanced neural network architectures, such as attention-based mechanisms, have shown success in capturing long-term dependencies and complex interactions among features in click prediction tasks. Models like DeepFM and Wide & Deep Learning, designed by Google, are also widely applied for ad click predictions.

4. Data Imbalance and Handling: Since the number of actual clicks in an ad click dataset is usually much lower than non-clicks, handling data imbalance is a crucial step. Techniques like under-sampling, over-sampling, or using synthetic data generation (e.g., SMOTE) are common approaches.

5. Data Privacy and Ethics: Given that ad click prediction often involves user-specific data, privacy-preserving techniques are critical. Techniques like federated learning or differential privacy are increasingly applied.

Dataset URL - https://www.kaggle.com/datasets/marius2303/ad-click-prediction-dataset

MADE BY:

ABHISHEK JAIN – E23CSEU0265

ANSH PANWAR – E23CSEU0256

DEV THUKRAL – E23CSEU0258