GrowthLink

Empowering Students, Elevating Careers

Machine Learning Assignment

Internship Opportunity Details:

Duration: 1 month
Mode: Remote
Work Hours: Part-time, Flexible (Easily manageable alongside college)
Stipend: ₹5000

Instructions

To proceed further, you need to complete any **one** of the five tasks listed below. Follow these guidelines to ensure your submission meets our requirements:

GitHub Repository Requirements:

- The task should have its own dedicated GitHub repository.
- Include a README file describing:
- Task objectives
    - Steps to run your project
    - Ensure clean, well-structured, and well-commented code.

Evaluation Criteria:

- Functionality: How well the task is implemented and works.
- Code Quality: Structure, readability, and efficiency.
- Innovation & Creativity: Unique features or optimizations.
- Documentation: Clarity in explaining the implementation.

Task 1: Movie Genre Classification

- The objective is to develop a model that classifies movies into genres.
- Requires working with text-based movie plot descriptions, which need to be processed into a format suitable for machine learning models.
- Involves applying appropriate techniques to convert raw text into numerical vectors.
- Various classifiers should be explored to determine the best fit for classification.
- Expected Outcome: A well-trained model capable of accurately classifying movies into genres based on textual descriptions, along with insights into feature importance and misclassifications.

DATASET

Task 2: Credit Card Fraud Detection

- The objective is to build a machine learning model that can distinguish fraudulent transactions from legitimate ones.
- The dataset contains transaction details such as amount, user ID, merchant information, timestamps, etc which must be carefully preprocessed.
- Identifying key transactional attributes that distinguish fraudulent behavior.
- Expected Outcome: A fraud detection system that minimizes false positives while maximizing fraud detection accuracy, along with an explanation of misclassifications.

DATASET

Task 3: Customer Churn Prediction

- The objective is to predict whether a customer will discontinue a subscription-based service.
- Historical customer data should be analyzed, considering factors like usage patterns, demographic details, subscription duration, etc.
- Customer data may have missing values, appropriate imputation strategies must be applied.
- Expected Outcome: A model that effectively identifies customers likely to leave, along with insights into the most significant factors contributing to churn.

DATASET

Task 4: Spam SMS Detection

- The objective is to develop an SMS classification model that identifies spam messages.
- The dataset consists of labeled messages categorized as spam or non-spam.
- Expected Outcome: A robust model that accurately distinguishes between spam and legitimate messages, with well-documented preprocessing and classification approaches.

DATASET

Task 5: Handwritten Text Generation

- The objective is to generate realistic handwritten-style text using deep learning models.
- Adjusting learning rates, sequence lengths, and model layers to optimize text generation quality.
- The generated text is assessed based on its coherence, readability, and similarity to actual handwritten text.
- Expected Outcome: A generative model capable of producing handwritten-style text that mimics human writing patterns, with detailed documentation on model training and evaluation.

DATASET

Submission Guidelines

How to Submit?

1. Upload your completed projects to GitHub.
2. Fill out the Submission Form: **form**
3. Ensure your GitHub repositories are public.
4. GrowthLink team will evaluate your submission within 24 hours.
5. The **intership completion certificate** and **LOR** whould be delievered to you within 24 hours of submission.
6. Deadline is mentioned in the email itself.

For Any Help or Queries:

Email: **help.growthlink@gmail.com**
Website: **GrowthLink**