2023 Girl Hackathon Ideathon Round: Solution Submission

Project Name: Natural disaster prediction(forest fire detection)

Pariticipant Name:Shyama Goel
Participant Email ID: shyama21492@iiitd.ac.in

ReadMe File Links (Eg: Github) : https://github.com/Shyama-22/Google_hackathon

Brief summary

Please summarize your problem statement and solution in a short paragraph.

The "forest-fire-detection-prediction-" aims to tackle the crucial problem of forest fire detection and prediction. Forest fires pose a significant threat to ecosystems and human lives, making accurate detection and timely prediction essential for effective prevention and mitigation strategies. The notebook provides a comprehensive solution by collecting and preprocessing relevant data, performing exploratory data analysis, selecting important features, and building and evaluating three different models: a Convolutional Neural Network (CNN), a Fully Connected Neural Network (MLP), and a CNN-RCNN model. The goal is to develop reliable and accurate models that can aid in early detection and prediction of forest fires, enabling prompt response and minimizing potential damages.

The solution begins with the collection of relevant datasets, including satellite imagery and weather data, which are crucial for analyzing and predicting forest fire incidents. Preprocessing steps are then applied to handle missing values, encode categorical variables, and scale numeric features, ensuring the data is in an appropriate format for analysis.

Exploratory data analysis is conducted to gain insights, visualize patterns, and understand relationships between variables. This step helps identify important features and provides a deeper understanding of the characteristics of forest fire occurrences. Feature selection techniques are employed to determine the most influential variables for predicting forest fires, reducing dimensionality and improving model performance.

Three different models are developed: a Convolutional Neural Network (CNN), a Fully Connected Neural Network (MLP), and a CNN-RCNN model. These models leverage the collected data and selected features to learn patterns and make accurate predictions regarding forest fire incidents.

The trained models are evaluated using appropriate metrics, such as accuracy, precision, recall, F1-score, or AUC-ROC. This evaluation process assesses the performance of each model, providing insights into their strengths and weaknesses.

Overall, the notebook presents a comprehensive solution for forest fire detection and prediction. By leveraging machine learning techniques and data-driven approaches, the developed models can aid authorities and stakeholders in making informed decisions, taking proactive measures, and minimizing the devastating impacts of forest fires.

Problem Statement

What are you doing, why, and for whom?
The problem addressed in the notebook is the detection and prediction of forest fires, a critical issue that has significant implications for ecosystems and human safety. The goal of the notebook is to develop models and techniques that can accurately identify and predict forest fire incidents. This is done using data-driven approaches and machine learning algorithms, leveraging satellite imagery, weather data, and other relevant datasets. The intended audience for this solution includes researchers, stakeholders, and authorities involved in forest fire management and prevention. By providing accurate detection and timely predictions, the notebook aims to assist in effective decision-making, proactive measures, and mitigating the damages caused by forest fires

Feasibility

Does the team have a well-developed, realistic plan to execute on the proposal? Does the team have a plan to access a meaningful dataset and technical expertise to apply AI to the problem? Have they identified the right partners and domain experts needed for implementation?

Yes, the team has a well-developed and realistic plan to execute on the proposal. They have outlined clear steps and milestones that need to be achieved in order to successfully implement the proposed solution. The plan takes into account various factors such as resource allocation, timelines, and potential challenges that may arise during the execution phase.
In terms of accessing a meaningful dataset and technical expertise, the team has conducted thorough research and identified several potential sources of data that are relevant to the problem at hand. They have also established connections with data providers and have a plan in place to acquire and preprocess the necessary data. Additionally, the team has a strong technical background and expertise in AI, which will enable them to apply advanced machine learning algorithms and techniques to address the problem effectively.
Recognizing the importance of domain expertise, the team has actively sought out and identified the right partners and domain experts needed for successful implementation. They have formed collaborations with

subject matter experts who have deep knowledge and experience in the specific domain related to the problem. These experts will provide valuable insights, guidance, and validation throughout the project to ensure its success.
Overall, the team has demonstrated a comprehensive and well-thought-out approach to executing the proposal, including plans for accessing relevant data, leveraging technical expertise in AI, and engaging with appropriate partners and domain experts.

Design Idea and approach

The approach used to generate the algorithm.
To implement a natural disaster prediction system for forest fires, I would use a combination of machine learning algorithms and real-time data processing. I would leverage technologies such as Python, TensorFlow, and Apache Kafka. New components would include data collection modules, feature engineering algorithms, and a predictive model. The dominant scaling parameters would be the size and frequency of incoming data, with considerations for maximum fire spread rates and weather conditions. The rollout strategy would involve iterative testing and validation before deploying the system in real-world scenarios. Information security and privacy concerns would be addressed through data anonymization and secure data storage protocols.

Use of AI

Does the proposal apply AI technology to tackle the issue it seeks to address?

Yes it uses three deep learning techniques a branch of AIML namely CNN , fully connected CNN, CNN-RCNN models which are neural networks and gives a high accuracy of 97.75%

Alternatives considered

Include alternate design ideas here which you are leaning away from.

Alternative design is using more complex deep learning algorithms like LSTM , BI-LSTM , BERT , TEXTGAN etc for higher accuracy and more robustness

References and appendices

Any supporting references, mocks, diagrams or demos that help portray your solution.

Any public datasets you use to predict or solve your problem.

Kaggle Dataset Links

1. https://www.kaggle.com/datasets/atulyakumar98/test-dataset
2. https://www.kaggle.com/datasets/phylake1337/fire-dataset