Data Collection and Preprocessing Phase

Date	03-10-2024
Team ID	LTVIP2024TMID24897
Project Title	Flight delays prediction using ML
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

A data collection plan for flight delays prediction involves gathering relevant data from various sources to identify the predict for flight delay. Raw data sources can include regression data, algoritms results. Additionally, data from wearable devices, mobile apps, and passenger-reported outcomes can also be used. The data collection plan should ensure that the data is accurate, complete, and relevant to the research question.

Data Collection Plan Template

Section	Description
Project Overview	The objective of this machine learning project is to identify the persons who have flight delays based on the prediction records.
Data Collection Plan	The data will be searched from the kaggle and other websites and world airlines websites.
Raw Data Sources Identified	The data will be collected from the kaggle and other websites and world airlines websites.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Flight airlines	The airlines details is valuable collections of data that includes 20 delays predict	kaggle datasets download -d usdot/flight-delays	CSV	1.2 MB	Public
Dataset 2	This dataset also contains the same common values as	https://www.kaggle.com/code/fabiendaniel/predicting-flight-delays-tutorial?scriptVersionId=1566193&cellId=1	Cvs	200 MB	public