Data Collection and Preprocessing Phase

Date	8th July 2024
Team ID	SWTID1720449665
Project Title	Predicting The Energy Output Of Wind Turbine Based On Weather Condition
Maximum Marks	2 Marks

Data Collection Plan Template

Section	Description
Project Overview	This machine learning project aims to predict wind turbine energy output based on weather data inputs, leveraging regression or machine learning models to optimize operational efficiency and ensure stable energy supply for wind farm operators.
Data Collection Plan	The data for this project can be collected from meteorological agencies, weather APIs (such as OpenWeatherMap or WeatherStack), historical weather databases, and potentially from sensors installed on wind turbines themselves for real-time measurements. We took the data from Kaggle.
Raw Data Sources Identified	Historical Weather Databases: Databases storing archived weather data spanning several years.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle	1. LV ActivePower (kW): The power generated by the turbine for that moment 2. Wind Speed (m/s): The wind speed at the hub height of the turbine (the wind speed that turbine use for electricity generation) 3.Theoretical_Power_Curve (KWh): The theoretical power values that the turbine generates with that wind speed which is given by the turbine manufacturer 4. Wind Direction (°): The wind direction at the hub height of the turbine (wind turbines turn to this direction automaticly)	https://drive.google.com/file/d/1s8DCU3CdEkYEtcxTrpv1LE16WaHgqj38/view	CSV	3.2 MB	Public