CS 451 - Data Mining

Flight Delay Prediction using Scalable Data Mining

Group Members –

Aakash Sinha – U101114FCS214

Rohan Shrivastava – U101114FCS118

Chalamala Krishnapriya – U101114FCS189

Work Plan for the Project –

Estimated Deadline	Tasks
Mid Sem - I	• Choosing a suitable Dataset • Understanding the data • Data Visualization
Mid Sem – II	• Data Sanitization • Attribute Selection • Sub-setting of Data • Replacing Missing Values • Removing Outliers • Re-Creating Variables
Comprehensive	• Creating Training and Testing Dataset • Applying different Models like Naïve Bayes, KNN, Decision Trees • Evaluation of all the algorithms • Deployment • Testing • Data Visualization

Roles of Individual Members –

Name	Roles
Aakash Sinha	- Data Pre-Processing - Watson Analytics - Coding - WEKA Testing - Deployment of Models - UI/UX
Rohan Shrivastava	- Data-Preprocessing - Coding - Testing - Deployment of Models
Chalamala Krisnapriya	- Data Pre-Processing - Documentation - Quality Assurance