Criteria	dict_keys(['Model Implementation and Ensemble Method', 'Data Preprocessing and Feature Selection', 'Evaluation Metrics and Model Performance'])	dict_keys(['Model Implementation and Ensemble Method', 'Data Preprocessing and Feature Selection', 'Evaluation Metrics and Model Performance'])	dict_keys(['Model Implementation and Ensemble Method', 'Data Preprocessing and Feature Selection', 'Evaluation Metrics and Model Performance'])	dict_keys(['Model Implementation and Ensemble Method', 'Data Preprocessing and Feature Selection', 'Evaluation Metrics and Model Performance'])
Model Implementation and Ensemble Method	The ensemble model is well-implemented, and the ensemble method is correctly applied. All required models (Multi-Kernel SVM, Decision Tree Classifier, and Logistic Regression) are built from scratch using only pandas and numpy. The ensemble method is clearly explained, and the code is well-organized and readable.	The ensemble model is implemented, but there might be minor issues with the ensemble method. Most required models are built from scratch, but there might be some minor errors or omissions. The ensemble method is explained, but the code might not be as well-organized or readable.	The ensemble model is partially implemented, and the ensemble method is not clearly applied. Some required models are built from scratch, but there are noticeable errors or omissions. The ensemble method is not clearly explained, and the code is not well-organized or readable.	The ensemble model is not implemented, or the ensemble method is not applied. Required models are not built from scratch, or there are significant errors or omissions. The ensemble method is not explained, and the code is not readable.
Data Preprocessing and Feature Selection	Data preprocessing is thorough, and feature selection is well-justified. All required steps (handling missing data, filtering active flights, etc.) are correctly applied. New features are created, and the reasoning behind feature selection is clearly explained.	Data preprocessing is mostly correct, and feature selection is justified. Most required steps are applied, but there might be minor issues. New features are created, but the reasoning behind feature selection might not be as clear.	Data preprocessing is partially correct, and feature selection is not well-justified. Some required steps are applied, but there are noticeable errors or omissions. New features are not created, or the reasoning behind feature selection is not clear.	Data preprocessing is not thorough, or feature selection is not justified. Required steps are not applied, or there are significant errors or omissions. New features are not created, and the reasoning behind feature selection is not explained.
Evaluation Metrics and Model Performance	Evaluation metrics (accuracy and F1-score) are correctly applied, and model performance is thoroughly evaluated. Cross-validation is performed, and the results are clearly explained.	Evaluation metrics are mostly correctly applied, and model performance is evaluated. Cross-validation is performed, but the results might not be as clear.	Evaluation metrics are partially correctly applied, and model performance is not thoroughly evaluated. Cross-validation is not performed, or the results are not clear.	Evaluation metrics are not correctly applied, or model performance is not evaluated. Cross-validation is not performed, and the results are not explained.

Code Quality, Readability, and Report	Code is well-organized, readable, and follows best practices. The report is clear, concise, and thoroughly explains the approach, results, and conclusions.	Code is mostly well-organized and readable, but might have some minor issues. The report is clear, but might not be as concise or thorough.	Code is partially well-organized and readable, but has noticeable issues. The report is not clear, concise, or thorough.	Code is not well-organized, readable, or follows best practices. The report is not clear, concise, or thorough.