

Name	Task	Database	Database URL
Abdul Malek Adib Mukhriz	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, ...). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Iris (UCI)	https://archive.ics.uci.edu/ml/datasets/iris
Abu Lail Habib Maher	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Titanic (Kaggle)	https://www.kaggle.com/competitions/titanic
Adorjáni Jonathán	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or ...). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	House Prices – Ames (Kaggle)	https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data
Alhijazin Katrin Sakher Ibrahim	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Wine Quality (UCI)	https://archive.ics.uci.edu/ml/datasets/wine+quality
Almásí Júlia	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Handle extreme class imbalance using class-weight adjustment. - Compare model performance before and after balancing. 	Credit Card Fraud (Kaggle)	https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
Amangeldin Arlan	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Adult Census Income (UCI)	https://archive.ics.uci.edu/ml/datasets/adult
Amirkhan Nurgul	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Pima Indians Diabetes (Kaggle)	https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database
Balekenov Zhanbolat	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use weather and temporal features. - Analyze seasonal patterns and evaluate models on subsets (e.g., winter vs summer). 	Bike Sharing (UCI)	https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset
Bartal Dávid	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Perform data augmentation and normalization. 	MNIST	https://archive.ics.uci.edu/dataset/683/mnist+database+of+handwritten+digits

Name	Task	Database	Database URL
Bernáth Máté Levente	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Define time-window features (e.g., hour, weekday, season). - Perform spatial feature engineering (pickup/dropoff clustering or ZIP-code aggregation). - Optionally model separate subsets (weekdays vs weekends). 	NYC Taxi (TLC)	https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page
Bilkevicius Motiejus	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use sliding windows and create multi-step forecasting targets. - Experiment with resampling (hourly/daily averages). 	ElectricityLoadDiagrams20112014	https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014
Bobkov Aleksei	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Iris (UCI)	https://archive.ics.uci.edu/ml/datasets/iris
Chang Jia Qian	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Titanic (Kaggle)	https://www.kaggle.com/competitions/titanic
Cory Jonathan David	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	House Prices – Ames (Kaggle)	https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data
Egorov Andrey	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Wine Quality (UCI)	https://archive.ics.uci.edu/ml/datasets/wine+quality
Fodor Kornél Máté	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Handle extreme class imbalance using SMOTE, undersampling, or class-weight adjustment. - Compare model performance before and after balancing. 	Credit Card Fraud (Kaggle)	https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
Gong Jiayu	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Adult Census Income (UCI)	https://archive.ics.uci.edu/ml/datasets/adult
Hennig Luca Silvio	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Pima Indians Diabetes (Kaggle)	https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database

Name	Task	Database	Database URL
Khudyeva Aydan	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use weather and temporal features. - Analyze seasonal patterns and evaluate models on subsets (e.g., winter vs summer). 	Bike Sharing (UCI)	https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset
Kiss Boróka	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Perform data augmentation and normalization. 	MNIST	https://archive.ics.uci.edu/dataset/683/mnist+database+of+handwritten+digits
Kossuth Hugó Ádám	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Define time-window features (e.g., hour, weekday, season). - Perform spatial feature engineering (pickup/dropoff clustering or ZIP-code aggregation). - Optionally model separate subsets (weekdays vs weekends). 	NYC Taxi (TLC)	https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page
Margitics László	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use sliding windows and create multi-step forecasting targets. - Experiment with resampling (hourly/daily averages). 	ElectricityLoadDiagrams20112014	https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014
Mátay-Csepeli Gábor András	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Iris (UCI)	https://archive.ics.uci.edu/ml/datasets/iris
Matesz Réka	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Titanic (Kaggle)	https://www.kaggle.com/competitions/titanic
Medvedev Matvei	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	House Prices – Ames (Kaggle)	https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data
Medvegy Gábor	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Wine Quality (UCI)	https://archive.ics.uci.edu/ml/datasets/wine+quality
Miha Anna	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Handle extreme class imbalance using SMOTE, undersampling, or class-weight adjustment. - Compare model performance before and after balancing. 	Credit Card Fraud (Kaggle)	https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Name	Task	Database	Database URL
Nagyhegyesi Luca	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Adult Census Income (UCI)	https://archive.ics.uci.edu/ml/datasets/adult
Nagy Mihály Gyula	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Pima Indians Diabetes (Kaggle)	https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database
Nagy Tamás	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use weather and temporal features. - Analyze seasonal patterns and evaluate models on subsets (e.g., winter vs summer). 	Bike Sharing (UCI)	https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset
Nguyen Hong Gia Bao	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Implement a Convolutional Neural Network (CNN) as the main model. - Compare CNN performance with a classical ML model such as SVM. - Perform data augmentation and normalization. 	MNIST	https://archive.ics.uci.edu/dataset/683/mnist+database+of+handwritten+digits
Nguyen Thanh Trung	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Define time-window features (e.g., hour, weekday, season). - Perform spatial feature engineering (pickup/dropoff clustering or ZIP-code aggregation). - Optionally model separate subsets (weekdays vs weekends). 	NYC Taxi (TLC)	https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page
Onder Sára	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use sliding windows and create multi-step forecasting targets. - Experiment with resampling (hourly/daily averages). 	ElectricityLoadDiagrams20112014	https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014
Péch Kornél	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Iris (UCI)	https://archive.ics.uci.edu/ml/datasets/iris
Schmiedt Lilla	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Titanic (Kaggle)	https://www.kaggle.com/competitions/titanic
Sólyom Lívia	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	House Prices – Ames (Kaggle)	https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data

Name	Task	Database	Database URL
Somlai Dóra	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Wine Quality (UCI)	https://archive.ics.uci.edu/ml/datasets/wine+quality
Szabó Bence	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Handle extreme class imbalance using SMOTE, undersampling, or class-weight adjustment. - Compare model performance before and after balancing. 	Credit Card Fraud (Kaggle)	https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
Szugiczki Flóra	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Adult Census Income (UCI)	https://archive.ics.uci.edu/ml/datasets/adult
Tóth Borbála Vivien	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. 	Pima Indians Diabetes (Kaggle)	https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database
Török Zsófia	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use weather and temporal features. - Analyze seasonal patterns and evaluate models on subsets (e.g., winter vs summer). 	Bike Sharing (UCI)	https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset
Tulit Zsolt	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Perform data augmentation and normalization. 	MNIST	https://archive.ics.uci.edu/dataset/683/mnist+database+of+handwritten+digits
Tursynbay Zhanlyva	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Define time-window features (e.g., hour, weekday, season). - Perform spatial feature engineering (pickup/dropoff clustering or ZIP-code aggregation). - Optionally model separate subsets (weekdays vs weekends). 	NYC Taxi (TLC)	https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page
Tussipbek Madi	<p>Project Type: Regression.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Linear Regression and a Decision Tree Regressor. - Train at least one advanced model (e.g., Boosting Regressor, Random Forest Regressor). - Provide evaluation using: MAE, RMSE, R² score, and residual-analysis plots. <p>Additional Requirements:</p> <ul style="list-style-type: none"> - Use sliding windows and create multi-step forecasting targets. - Experiment with resampling (hourly/daily averages). 	ElectricityLoadDiagrams20112014	https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014
Zádor Levente	<p>Project Type: Classification.</p> <p>Your task is to build a complete machine-learning pipeline using the selected dataset. This should include exploratory data analysis, data cleaning, feature engineering, model training, hyperparameter tuning, and evaluation. Provide clear visualizations that support your decisions.</p> <ul style="list-style-type: none"> - Train at least two baseline models, including Logistic Regression and a Decision Tree. - Train at least one advanced model (e.g., Random Forest, Boosting, or SVM). - Provide evaluation using: confusion matrix, classification report, ROC curve. 	Iris (UCI)	https://archive.ics.uci.edu/ml/datasets/iris