**Oil Spill Dataset Analysis and Prediction**

Using the Oil Spill Dataset, this Model analyzes and predicts oil spills in satellite images. The dataset contains satellite images divided into patches, which have been processed using computer vision algorithms to extract features describing their contents. The goal is to predict whether a given patch contains an oil spill or not.

**Dataset Information**

The Oil Spill Dataset consists of satellite images split into patches. The dataset has the following characteristics:

* Total Columns: 50
* Target Column: 'target'
* Classes:
  + Non-Spill: Negative case (majority class)
  + Oil Spill: Positive case (minority class)

**System Requirements**

To successfully run the code and perform the analysis, your system should meet the following requirements:

* Python (version 3.6 or higher)
* Jupyter Notebook or any Python IDE (e.g., Anaconda, PyCharm)
* Required Python libraries: pandas, numpy, matplotlib, seaborn, scikit-learn

**Tasks**

1. **Data Cleaning and Pre-Processing:**
   * Load the dataset using pandas
   * Drop rows with missing values
   * Perform one-hot encoding for categorical variables
   * Save the cleaned dataset to a new CSV file
2. **Deriving Insights from the Dataset:**
   * Check column names and summary statistics
   * Explore data types of each column
   * Calculate mean, median, and standard deviation of specific columns
   * Count the occurrences of a specific value in the target column
   * Visualize data using histograms and bar plots
   * Analyze correlations using a correlation matrix and heatmap
3. **Applying Machine Learning Techniques:**
   * Split the data into features and target
   * Split the data into training and testing sets
   * Scale the data using StandardScaler
   * Implement individual classifiers (Decision Tree, Random Forest, Logistic Regression, KNN)
   * Apply BaggingClassifier, RandomForestClassifier, and VotingClassifier (Ensemble)
   * Evaluate the models using accuracy score
   * Save the best model using joblib
4. **Loading the Saved Model and Making Predictions:**
   * Create a new dataset by randomly selecting 20 data points from the original dataset
   * Load the saved best model
   * Set the feature names explicitly
   * Make predictions on the new dataset
   * Print the predictions

**Repository Structure**

* **data/**: Folder containing the Oil Spill Dataset and the randomly generated dataset (if applicable)
* **notebooks/**: Folder containing Jupyter Notebooks or Python scripts for data cleaning, analysis, and modeling
* **models/**: Folder to store the saved best model
* **results/**: Folder to store evaluation results, outputs, and the cleaned dataset

**Instructions**

1. Download the Oil Spill Dataset and place it in the **data/** folder.
2. Run the provided code in the Jupyter Notebook or Python script step by step to perform data cleaning, analysis, and modeling.
3. Modify the code or add new code as needed to meet the assignment requirements.
4. Save the cleaned dataset to the **results/** folder as **cleaned\_dataset.csv**.
5. Update the **README.md** file with any additional instructions or explanations, if necessary.
6. Run the code to save the best model to the **models/** folder as **best\_model\_Oil\_Spill.pkl**.
7. Create a new dataset by randomly selecting 20 data points from the original dataset.
8. Save the randomly generated dataset to the **data/** folder as **random\_dataset.csv**.
9. Update the `results