

# Project

Smishing Detection in Telecommunications: The Efficacy of Textual Feature Integration in Machine Learning Models

## Prerequisites

Before you begin, ensure you have met the following requirements:

- You have installed the latest version of Python.

## Installation

To install the project, follow these steps:

Linux and macOS:

```
python3 -m venv .venv
source .venv/bin/activate
pip install -r requirements.txt
```

Windows:

```
python -m venv .venv
.\.venv\Scripts\activate
pip install -r requirements.txt
```

### Using the Project

To use the project, follow these steps:

Before running the notebooks, ensure you have the following datasets in the `data` directory.

- `merged.csv`
- `MOCK_DATA.csv`
- `MOCK_DATA-2.csv`
- `MOCK_DATA-3.csv`

- `MOCK_DATA-4.csv`
- `MOCK_DATA-5.csv`
- `spam.csv`

To run the notebooks correctly, it has to be run in the following order:

- `data_merge.ipynb`
- `eda.ipynb`
- `model_training_standard.ipynb`
- `model_training_nlp.ipynb`