

# MLOps for Document Classification Models

## Data Preparation:

1. **Data Collection:** Select or download data for training a document classification model. The data should include the text of documents and corresponding labels.
2. **Data Processing:** Use NLP libraries such as spaCy, NLTK, or scikit-learn to process the data. This includes tasks like tokenization, removing stop words, and text vectorization.

## Model Training:

1. **Use scikit-learn or Other ML Libraries:** Train a document classification model using scikit-learn or other ML libraries on the processed data.
2. **Evaluate Model Performance:** Evaluate the performance of the model on a test dataset.

## Model Packaging with BentoML:

1. **Install BentoML:** Start by installing the BentoML framework, designed for packaging and serving ML models in production.
2. **Define and Save the Model:** Use BentoML to define the trained model and package it into a container. This container includes the model, dependencies, and necessary configurations.

## Configuring BentoML for JSON and Multi-Part Serialization:

1. **Specify Data Serialization:** Configure BentoML to handle data serialization, with a focus on JSON and Multi-Part serialization. This enables efficient and easy transmission of data for model predictions.

## Model Serving:

1. **Configure BentoML Server:** Set up a BentoML server to serve the packaged model. The server can run either locally or on a cloud-based server.
2. **Expose an API Endpoint:** Create an API endpoint that allows Python clients to submit documents for classification. Ensure that the endpoint supports data serialization in JSON and Multi-Part formats.

### Model Testing:

1. **Develop a Python Client:** Develop a Python client that sends documents to the model-serving endpoint for classification, using JSON and Multi-Part serialization.
2. **Verify Model Predictions:** Ensure that the model returns correct predictions and handles data serialization appropriately.

### Documentation and Presentation:

1. **Document the Entire Process:** Document the entire process, including data processing, model training, packaging with BentoML, and serialization configurations.
2. **Present the Project:** Present the project and the developed solution to colleagues and instructors for evaluation.

### Suggested Tech Stack:

- **Python**
- **NLP Libraries:** spaCy, NLTK, etc.
- **ML Libraries for Model Training:** scikit-learn or other ML libraries
- **BentoML for Model Packaging and Serving:** with support for JSON and Multi-Part serialization
- **Python Client for Testing and Usage**