### ML 23/24-10 Multi-Sequence Learning with language semantic

# Multi sequence Learning

## 1. Introduction

### 1.1 Purpose

The purpose of this document is to provide comprehensive documentation for the Multi sequence Learning Experiment. This experiment aims to demonstrate the learning and prediction capabilities of a model when dealing with sequences of data, specifically applied to language semantics.

### 1.2 Scope

This documentation covers the implementation details, usage, and key features of the Multi sequence Learning Experiment. It also provides guidance on modifying and extending the existing codebase.

## 2. Getting Started

### 2.1 Prerequisites

- .NET Framework installed

- Basic understanding of C# programming language

### 2.2 Installation

Clone or download the Multi sequence Learning repository to your local machine.

```bash

git clone https://github.com/your\_username/multisequence-learning.git

```

## 3. Experiment Overview

### 3.1 Experiment Objective

The primary objective of this experiment is to learn and predict sequences based on the provided training data. The experiment utilizes a model that can be trained on sequences and then used to predict the continuation of given input text.

### 3.2 Existing Code

The existing code is implemented in the `MultisequenceLearning.cs` file. It includes a sample `RunMultiSequenceLearningExperiment` method for learning and predicting sequences.

## 4. Modifications and Enhancements

### 4.1 New Method: RunLanguageSemanticExperiment

A new method, `RunLanguageSemanticExperiment`, has been introduced to enhance the existing experiment. This method enables automatic reading of sequences from a file, splitting data into training and testing sets, and calculating prediction accuracy.

### 4.2 Loading Sequences from File

The `LoadSequencesFromFile` method reads text data from a file, removes control characters, and converts characters into ASCII codes. It returns a list of integer arrays representing sequences.

### 4.3 Training and Testing Data Split

The training and testing data are split with 90% used for training and 10% for testing. This ensures that the model is trained on a substantial portion of the data and tested on a separate set.

### 4.4 Model Training

The `Model` class includes a `Train` method for training the model with the provided training data.

### 4.5 Accuracy Calculation

The `CalculateAccuracy` method is used to calculate the accuracy of the model on the testing data. Binary cross-entropy or other suitable metrics can be implemented based on the user's requirements.

### 4.6 Prediction (Inference)

The `Predict` method in the `Model` class allows users to enter text, and the model generates and returns predicted text based on the input.

## 5. Usage

### 5.1 Running the Experiment

To run the experiment, execute the `Main` method in the `MultisequenceLearning.cs` file. This will invoke the `RunLanguageSemanticExperiment` method.

### 5.2 Input Files

- `trainingFilePath`: Path to the file containing the training text data.

- `testFilePath`: Path to the file containing the testing text data.

## 6. Conclusion

The Multisequence Learning Experiment provides a foundation for understanding and implementing sequence learning and prediction. Users can modify and extend the code to suit specific applications and experiment with different types of data.