**MGSC 695 – Assignment 2**

Adv Topics in Mgmt Science

Konstantin Volodin – 261083570

Table of Contents

[1. Introduction 1](#_Toc139223005)

[2. Data Review 1](#_Toc139223006)

[3. Network Architecture 1](#_Toc139223007)

[4. Final Model Results 1](#_Toc139223008)

## Introduction

The objective of this project is to create a recurrent neural network (RNN) capable of generating coherent text. Our approach involves employing a comprehensive machine learning pipeline, encompassing data exploration, model development using PyTorch Lightning, and rigorous model evaluation.

## Data Review

For this project, we utilize a dataset consisting of Shakespeare's writings, specifically the play "Macbeth," which contains a total of 2586 lines of text. Our goal is to predict individual characters based on a list of input characters. In this setup, we use 100 input characters and 1 output character, resulting in a total of 94,077 combinations for training. Additionally, the characters are label encoded, and input characters are scaled to a range of 0-1. Due to time constraints, we retain a portion of the data for training (32,000 samples) and validation (1,600 samples).

## Network Architecture

We employ a simple LSTM network with multiple layers and a single dense layer to convert the outputs to the appropriate format. The LSTM network is chosen to mitigate the issue of vanishing or exploding gradients over time. The specific configuration of the LSTM portion is as follows: 64-layer size, 1 layer, and a dropout rate of 0%. The model is trained for 25 epochs and evaluated thereafter. While the model improves over time (see Figures 1 and 2), the improvement is slow, suggesting more epochs or an improved architecture should be used (unfortunately for this assignment I didn’t have the time to properly experiment with it).

## Final Model Results

It is important to note that spaces and punctuation marks are counted as characters in this context. The final model is trained for 25 epochs, achieving a validation accuracy of approximately 36.4%. To assess the model's text generation capabilities, we provide the input text, "The luminous moon gr" and generate text for 20 characters. The resulting output text is as follows: "The luminous moon great the the the the ". While the model does predict some characters, in a different patter, it is not predicting well enough requiring further improvements.

Figure 1: Train Accuracy over Time

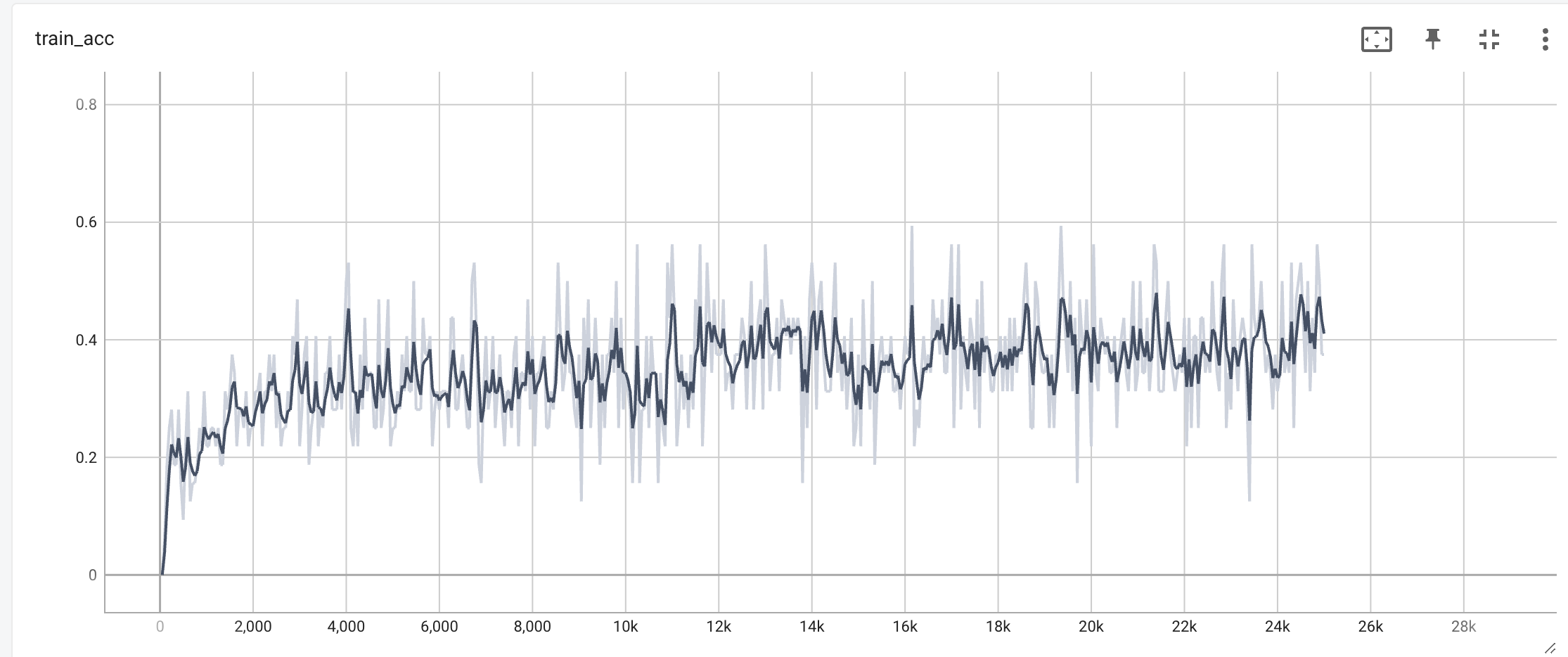


Figure 2: Validation Accuracy over Time

A picture containing line, plot

Description automatically generated