ADTA 5560.701
Recurrent Neural Networks for Sequence Data

Thuan L Nguyen, PhD

Midterm Assessment

1. Overview

The midterm covers all the topics that have been discussed in the first half of the course. The materials in any format, including in-class discussion, should be considered and used for the midterm. Additionally, the student can use any other source of information that he/she can gather, providing it is relevant and supporting the student's answers.

The student is required to create an MS Word document named "**ADTA_5560_midterm.docx**" that will contain all his/her midterm work, except for the Python coding.

IMPORTANT NOTES:
--) *--) If an MS Word document is specified as the required format of the submitted document, the student should **back up** the MS Word document by saving it as a PDF file before submitting it.*

IMPORTANT NOTES:
--) If an MS Word document is specified as the required format of the submitted document, the student should submit it, __not__ submit a PDF.

--) All the submission requirements are expected to be submitted in an MS Word document, except for Python code, or being specified otherwise.

--) For Python code in Jupyter Notebooks, the student is required to run the code and submit the Jupyter Notebook document that contains the results. The student should __not__ copy the results of Python code into the MS Word document.

IMPORTANT NOTES:
--) When discussing a topic or answering a question, it is expected that the student has to provide adequate explanations and supporting details.

2. Datasets

The data set used in this midterm assessment will be generated by the student following the instructor's instructions.

3. PART I: Learning Process in Neural Networks (20 Points)

SUBMISSION REQUIREMENT PART I:

Discuss the learning process of a neural network like Feedforward Neural Network (FFNN).

4. PART II: Sequence Data and Memory (20 Points)

SUBMISSION REQUIREMENT PART II:

Question 2.1:
--) Discuss the special relationship between sequence data like language and memory

Question 2.2:
--) Discuss why the recurrent neural network is a good fit for processing sequence data like language

5. PART III: Simple RNN Cell and McCulloch-Pitts Model (20 Points)

SUBMISSION REQUIREMENT PART III:

Question 3.1:
--) Discuss the Simple RNN Cell showing that it is a version of the McCulloch-Pitts model that is implemented in a real artificial neural network.

Question 3.2:
--) Discuss and prove that the Simple RNN cell has computation power.

6. PART IV: Simple RNN with Sine Wave Data and Keras (20 Points)

TO-DO

Build, train, and **evaluate** a simple recurrent neural network (a complete simple RNN) that has **two layers**: a SimpleRNN and a fully-connected layer in Keras. The **SimpleRNN layer** has **64** neurons.

The training and testing the model are done on **sine wave data** belonging to a dataset generated by dividing the **range 0 … 50** into **768 data** points, including 0 and 50. All the tasks related to the project are done in **Python** using **Jupyter Notebook**.

IMPORTANT NOTES:
--) It is expected that the student knows how to count the layers of a deep neural network.

--) First, **draw a diagram** to describe the architecture of the neural network with all the layers using MS PowerPoint or Draw Tool in MS Word.

--) **Build** and **train** the model using the Keras sequential model in a Jupyter Notebook document.
--) **Run all the steps** of the project in the Jupyter Notebook document to get the results of each step.

IMPORTANT NOTES:
--) It is expected that the student completes all the steps, including those related to data preprocessing.

--) **Evaluate** the model on the test data.
--) **Write a report** on these results.

SUBMISSION REQUIREMENT PART IV

--) Add **one section** to discuss the design of the neural network into the MS Word document, "**ADTA_5560_midterm.docx,**" including the **diagram** of the neural network.

--) Add **another section** to include the report on the results of training and evaluating the model into the document: "**ADTA_5560_midterm.docx,**"

--) **Submit** all the **code** and the **results** of running the code in a **Jupyter Notebook document** (*.ipynb)

7. PART V: Redesign Simple RNN (20 Points)

TO-DO
Redesign the simple recurrent neural network by changing the number of neurons of the SimpleRNN layer. The student can choose the number he/she wants.

--) **Rebuild** and **retrain** the **new model** using the Keras in **another Jupyter Notebook** document.
--) **Run all the steps** of the project in the Jupyter Notebook document to get the results of each step.

--) **Evaluate** the **new model** on the test data.
--) **Write a report** on these results.

SUBMISSION REQUIREMENT PART V:

--) To discuss the new design of the simple recurrent neural network, add **one new** section into the MS DOCS document above: "**ADTA_5560_midterm.docx**". The discussion should include:
- The **diagram** of the redesigned neural network
- Discussion in detail of **how** the Simple RNN is redesigned
- Discussion in detail of **why** such a redesigned network can potentially produce improved performance, i.e., higher accuracy level.

--) Add **another** section to the MS DOCS document: "**ADTA_5560_midterm.docx**" to discuss the results obtained from the redesigned Simple RNN, especially comparing them with those from PART IV.

IMPORTANT NOTES:
--) With the assumption that the student uses the trial-and-error approach, it is __OK__ if the results of training and evaluating the redesigned neural network do not show any significant improvement in the network performance.

8. HOWTO Submit

Due date & time: 11:00 PM – Saturday 11/09/2024

The student is required to submit the midterm – all the documents: Microsoft Word and Jupyter Notebooks – as attachments to a UNT email that is sent to the instructor (Thuan.Nguyen@unt.edu).

The subject of the email: "ADTA 5560: Midterm Assessment – Submission."