

Dissertation Schedule and Tasks

Type

Task

ID

42

Subject

Create Initial Work Plan for Infrastructure as Code

Status

New

Assignee

Bruce Goldfeder

Priority

High

Description

This is the task for creating the initial plan for setting up my Data Science environment where I can begin experimentation for my dissertation using Transformer based Neural Nets.

Type

Task

ID

43

Subject

Exercise Dissertation Neural Net Infrastructure with a RAD based language leveraging a pre-trained BERT Neural Net capability

Status

New

Assignee

Bruce Goldfeder

Priority

Normal

Description

Type

Dissertation Schedule and Tasks

Task

ID

44

Subject

Create and provide all of the specifications for a repeatable and automated capability to recreate initial RAD based language BERT Neural Net capability

Status

New

Assignee

Bruce Goldfeder

Priority

Normal

Description

Type

Epic

ID

45

Subject

Learn in Depth Transformer Neural Net Architecture

Status

New

Assignee

Bruce Goldfeder

Priority

Normal

Description

This task is to master the Transformer based neural net architecture. Initial investigation will be to leverage BERT as a foundation to extend into novel and unique model designed for Cyber Bullying detection.

Type

Epic

Dissertation Schedule and Tasks

ID
46
Subject
Learn in Depth PyTorch high level language supporting neural net model development
Status
New
Assignee
Bruce Goldfeder
Priority
Normal

Description
The PyTorch python library for the generation of neural nets will serve as my initial higher level library for the creation of transformer based neural net models. Alternatives to this library are Tensor Flow 2.x and/or Keras which are developed and used by Google, the creators of the BERT transformer model upon which my experiments are planned to work upon as foundation.

Type
Epic

ID
47
Subject
Create experimental set of data for Cyber Bullying detection
Status
New
Assignee
Bruce Goldfeder
Priority
Normal

Description
The data set for my experiments within my dissertation will be chosen from prior work done in the field. I will put together a matrix for deciding which data set or sets to use and apply heuristics and quantifiable attributes

Dissertation Schedule and Tasks

for
deciding the best data set to use.

Type

Milestone

ID

48

Subject

Complete Infrastructure Setup, Experimental Design, and Foundational Understanding of Transformer Neural Net Development

Status

New

Assignee

Bruce Goldfeder

Priority

Normal

Description

The **Infrastructure Milestone** represents the capstone of this semesters research. The primary goals of the Fall

2020 semester will be realized at the completion of this milestone. These are the creation of fully scripted and

automated experiment infrastructure in a Cloud environment, the design for research experiments for assessing

performance and accuracy of Cyberbullying Detection models, and a full foundational knowledge with supporting code

examples for Transformer based Neural Nets using the PyTorch library for development. **Infrastructure as Code**

encompasses setting up repeatable infrastructure in virtual environments (e.g. AWS, GCP, Azure, Lambda GPU). This

will provide a repeatable set of common infrastructure that can be leveraged for repeated experiments to provide

comparable runs with neural nets of varying architectures and configurations. **Experimental Design** represents the

creation of a set of experiments to prove the hypothesis of increased performance and accuracy for the detection of

Cyberbullying using Transformer based neural net language models. Part of this epic will be the selection of the

Cyberbullying data set or sets and prior state of the art examples for comparison. **Foundational**

Dissertation Schedule and Tasks

Understanding of

Transformer Neural Net Development** requires mastering the concepts and linear algebra mathematics underpinning the

Transformer architecture as realized in the Google BERT architecture. Full understanding of the supporting library,

PyTorch for realizing this model in code is also a prerequisite for beginning the experimentation planned for Spring

2021 semester.