DEPARTMENT *of* COMPUTING
College of Business & Technology
EAST TENNESSEE STATE UNIVERSITY

CSCI 5260 – Artificial Intelligence

Semester Project – An Application of Artificial Intelligence

A Phased Project

This project is research-oriented and is fairly open-ended. The project will explore an application of AI to a particular area that interests you. The "area" would be a specific application of AI in the world and could be nearly anything—from gameplay/development to self-driving vehicles or any other domain. This project will be iterative, beginning with a topic proposal and culminating in a paper and class presentation.

Assignment Structure

Phase 1: Proposal → Phase 2: Approach → Phase 3: Results → Phase 4: Presentation

Phase 1: Proposal

Choose a general topic of interest to you. This project aims to conduct basic experimentation using Artificial Intelligence and then report the results. One area of interest might be "network intrusion detection," "network routing," "program interactions," "public health," "stock market data," "gaming," "image processing," "baseball statistics"—really, you can choose any topic area that interests you.

Do some basic research to locate a data set that matches your topic area. Many data sets are available online, via websites like www.data.gov and www.kaggle.com. The data set should be large, but not so large that analysis is impossible. If you choose a large data set, you might need to use only a subset of the data to complete this project.

After locating a set of data, propose an artificial intelligence mechanism that will help you analyze it. Part of this process will be discovering the best algorithms that would apply to a given data set. Your analysis could involve searching for particular features, classifying data, performing natural language processing, constraint satisfaction, planning, or even basic machine learning (e.g., clustering analysis, statistical machine learning, neural networks, or support vector machines). You may also find that a combination of these techniques (ensemble techniques) would prove useful.

If you are working on a thesis, you may choose your thesis topic area, but you may not choose the same artificial intelligence technique for analysis.

Phase 1 Specification:

1. Write a detailed description of the data set you have chosen. Be sure to indicate the source of the data set. Websites like Kaggle.com list additional sources for the data, so you should include this

information as well. You may also use data sets from academic literature, and those must be cited as well. If you are required to narrow your data set (i.e., to limit the number of records because the initial data set is too large), explain exactly how you plan to do so.

2. Write a description of the artificial intelligence mechanism or mechanisms you plan to apply to the data set. Be sure to justify why using the AI techniques would assist in your exploration of the data set in question.

3. Write a description of what you believe this application of AI to this data set will produce. I am not looking for a prediction of the results here. I simply want a paragraph that describes the *value* of this application.

This deliverable must be no more than two pages in length, using 12-point Times-Roman font. Use APA or IEEE citation style.

Phase 2: Approach

With the topic you proposed in Phase 1 (and the feedback I provided), develop a detailed approach for analyzing the data. One old but helpful paper about writing approaches is: http://kmh-lanl.hansonhub.com/pc-26-58-seisler.pdf (focus on the Purpose, Activities, and Output sections).

Phase 2 Specification:

Your approach should describe:

1. Data manipulation, if necessary.
   a. You might determine that some data elements need to be ignored. Be sure to list those, or, if that list is too large, list the data elements you will use in your analysis.
   b. If you use a machine learning technique, be sure to state how you divide the training and test sets.
   c. If you have to translate your data into another format to make it analyzable, you should describe that here.
2. Data analysis.
   a. Explain the step-by-step process you will take to perform the AI analysis of your data set. You must be very specific here, so that means you should have done some preliminary exploration and research before writing this section.
   b. Explain how you will collect and report the results. Again, be very specific about this. Ideally, you would have a method for visualizing the results that you locate. The visualizations often depend on the method of AI algorithm you use. Thus, you could have anything from a heatmap to bar and pie charts. If possible, you should show an example of how your reporting mechanism will look—although it does not need to use real results at this point.

Because of the nature of this section, I am not setting a page limit. However, approaches that are too short or lack details will not be viewed favorably. You may send your approach to me before turning it in, and I will make comments as necessary.

Phase 3: Results

After submitting your approach, perform the analysis of your data set and collect the results. **You are required to use Python to complete the project.**

Deliverable 3 Specification:

You should prepare an academic report, **using the provided template**, with the following sections.

- **Front Matter** – Include a title and a list of authors.
- **Introduction** – Include a description of the project and the problem you are trying to solve. Draw information from Phase 1 for this.
- **Approach** – Under this heading, give the detailed list of steps you took (use past tense) to complete this project. Draw information from Phase 2 for this.
- **Results** – Report your findings under this heading, using the reporting mechanisms you described in Phase 2 (unless I have instructed you to do otherwise).
- **Conclusions** – State what conclusions you draw from the results presented.

- **References** – List all references, academic or otherwise, used in this project.

You must submit all code and the data set used for this project.

PHASE 4: PRESENTATION

Record a presentation that presents your topic and the results. Develop visual aids (e.g., PowerPoint) that discuss your findings and conclusions. The general outline of the presentation should be the same as Phase 3, and should be presented so that everyone in the class can understand it. Note that you might need to spend a little time explaining how your AI approach works. Demonstrations, where possible, are encouraged. This presentation should last about 15 minutes. Teams who go significantly over this time will receive a grade deduction.

Your recorded presentation will be posted to the D2L website for all students to view.

GRADING

The final grade for the semester project will be based on the following components, with the following weights. Each deliverable will be graded based on the specifications above using a letter grade system. The letter grades will be converted to numeric values based on the grading scale in the syllabus.

Phase 1: Proposal	15%	Sunday, February 14, 2021
Phase 2: Approach	25%	Sunday, February 28, 2021
Phase 3: Results (Includes Code and Report)	40%	Sunday, April 11, 2021
Phase 4: Recorded Presentation	20%	Sunday, April 25, 2021