College of Engineering and Computer Science
Department of Computer Science
Course Code: CSC-215 (Section 02)
Course Title: Artificial Intelligence
Semester: Spring 2025

Instructor:	Dr. Bonaventure Chidube Molokwu
Assessment Title:	Project 01 (AI-based Project)
Assessment Code:	CSC-215-02-P01
Duration/Deadline:	11.59pm, Monday, April 28

Group Name:	**project(Echo)**
Experiential Study on Societal Factors (Socio-Economic, Socio-Political, Socio-Cultural, etc.) that influence Financial Worthiness in the United States	Walid Hakimzada Andrew Nguyen Alexander Oswalt Michael Robertson Karthik Reddy Thippareddy

1. **Project Overview:**
   Our client is demanding for a Machine Learning (ML) application that can be trained on data to acquire knowledge, and learn how to predict the (expected) financial-income value of individuals based on its acquired knowledge with reference to varying factors which border on socio-economic axis, socio-political axis, socio-cultural axis, demographic axis, etc. This ML application is intended to assist Economic Experts, Financial Experts, Policy Makers, Employers of Labor, etc. make informed data-driven and empirical decisions with respect to understanding and/or enhancing the quality of life for humans. Moreover, this ML application should be capable of processing the impact and influence of socio-economic, socio-political, socio-cultural, demographic factors, etc. with respect to predicting the financial worthiness of an individual within a given territory. Additionally, the overall study and experimentation in relation to this project MUST be documented and compiled into an authentic (research) article worthy of presentation at a reputable Computer Science venue. The Institute of Electrical and Electronics Engineers (IEEE) conference proceedings template is available online via: IEEE Conference LaTex Template

   The aim of this (Graduate) project is to create a fully functional AI-based application using modern Machine Learning (ML) and/or Deep Learning (DL) algorithms and methodologies. This project will showcase the students' skill-set, achievements, teamwork, etc.; thereby providing potential employers or clients with an overview of the students' expertise with reference to the domain of Artificial Intelligence (AI).

2. **Project Objectives:**
   At the end of this AI-based Project, each participant/student will:

   (a) Be capable of employing useful Software Process (or Software Development Life Cycle) models with respect to teamwork in Software Engineering.

(b) Gain hands-on experience with respect to Machine Learning (ML) frameworks, viz: Scikit-Learn, XGBoost, LightGBM, CatBoost, H2O.ai, etc.

(c) Be capable of working with Deep Learning (DL) frameworks, viz: TensorFlow, PyTorch, Keras, MXNet, JAX, etc.

(d) Gain hands-on experience with respect to the usage of Python libraries, namely: Pandas, NumPy, Matplotlib, SciPy, Seaborn, Plotly, etc.

(e) Be able to store and manipulate datasets using any popular File System (.txt, .rdf, .csv, .xlsx, etc.) and/or Database Management technologies.

(f) Create a trained, intuitive, effective, efficient, and user-friendly AI-based application.

(g) Use Google Colab, Jupyter Notebook, etc., for teamwork and/or version management.

3. **Project Requirements:**
In addition to other essential and functional aspects, this AI-based application MUST include the following aspects, viz:

(a) preprocessing(Data): Detailed procedural approach with respect to handling missing data-points, transforming the data into numerical representation, and other relevant data cleaning techniques applied to the dataset.

(b) engineering(Feature): Details of how Feature Engineering, with respect to Feature Extraction and Feature Selection, was implemented during the training of a ML algorithm. Alternatively, this may represent details of how Feature/Representation Learning was implemented with respect to training a DL algorithm.

(c) transformation(Data): Expatiate on what transformation approach(es) is/are applied towards transforming the data into representations digestible by ML and/or DL algorithms. Thus, this may include Feature Standardization, Target Standardization, Feature Normalization, Target Normalization, etc.

(d) training(Algorithm): Detailed report on the training and validation strategy (k-fold cross validation, etc.) employed towards fitting and validating the ML or DL algorithm prior to *out-of-sample* testing.

(e) evaluation(Performance): Highlight on what type of objective functions, fitness functions and/or cost functions, that were employed towards evaluating the performance of the ML or DL model during *out-of-sample* testing. Also, the real values ($\mathbb{R}$), associated with all the performance-evaluation metrics, should be carefully reported and analyzed for problems associated with the ML or DL model *overfitting* or *underfitting* on the training data.

(f) compilation(Research Article): The aforementioned experimentation and work carried out herein MUST be documented into an article with original idea(s), and this article should be worthy of publication at a reputable Computer Science venue. The following structure may serve as a guide towards composing your original article, viz:

(i) Introduction: This section should briefly introduce the topic and its significance, define the Problem Statement, explain the motivation for this topic and why it is important, identify the original/novel idea(s) that the work herein addresses, state the specific goals of this work, and briefly outline the structure of the article.

(ii) Review of Literature and Related Work: This section gives a comprehensive review of existing literature relevant to this work, discuss the limitations of existing work or solutions, etc.

(iii) Materials and Methodology: This section describes the datasets, preprocessing techniques, etc., used for the experiments and work herein; describes the methodology applied towards the experimentation and work; describe the algorithms or models used in this work; and provide details on how the experiment(s) were implemented (e.g. frameworks, programming languages, libraries, etc.).

(iv) Experiments and Results: This section describes the experimental setup, parameters, etc.; it should clearly present the results and findings from the experiments; and use figures, tables, charts, etc., to effectively present the results and findings.

(v) Discussion: This section analyze the aforementioned experiment results and discuss their significance as well as implications, and it should clearly highlight the limitations of the work herein.

(vi) Conclusion and Future Work: Clearly summarize the key findings and contributions of the work herein, briefly reiterate the significance of the work and experiments herein, and suggest potential future work or directions based on your findings herein.

(vii) Acknowledgments: This section should acknowledge any support from instructor, supervisor, or financial/technical/resource support from funding organization.

(viii) References: This section should contain an ordered list of cited/referenced projects, work, articles, etc., using the IEEE Conference LaTex Template referencing style.

4. **Tools and Technologies:**

(a) Machine Learning (ML) frameworks: Scikit-Learn, XGBoost, LightGBM, CatBoost, H2O.ai, etc.

(b) Deep Learning (DL) frameworks: TensorFlow, PyTorch, Keras, MXNet, JAX, etc.

(c) Python libraries: Pandas, NumPy, Matplotlib, SciPy, Seaborn, Plotly, etc.

(d) Version Control: Colab, Jupyter Notebook, etc.

5. **Deliverables:**

(a) Source Code: Upon completion, kindly compile all the source codes, executable and/or binary files, datasets, resources, etc., into a folder/directory titled: `project_{Group Name}` for submission via Canvas^{TM} LMS.

- This folder/directory should include a README.txt file with the instructions on how to run and view the AI-based solution on a local Personal Computer (Windows OS platform, preferably) or over the World Wide Web (WWW).
- Include necessary comments in each code file for improved readability.

(b) Research Article/Paper: A detailed report (*6 . . . 8* pages in IEEE Conference LaTex Template format) which addresses the requirements specified in Section 3. above.

6. **Submission Instructions:**

- Kindly compile ALL the aforementioned deliverables into a compressed (`.zip`) file, and submit on/before the stipulated submission deadline via Canvas^{TM} LMS.
- Late Submission: Each late submission will receive a deduction of 20%, with respect to the total possible (100) points, for each day after the submission deadline. Submissions more than **5 days** late will receive zero (0) points.

7. **Presentation and Team Cohesion:**
The presentation(s) due dates are stated in the Course Syllabus. Also, a more detailed presentation schedule will become available as we approach the due date.

Team Structure and Cohesion: Each member, with respect to each project (team), MUST assume one or more role(s) as stated below with reference to the development of each AI-based Project.

Table 1: User Roles with reference to AI-based Project

Role	**Maximum**	**Description**
Project Manager	1	Responsible for the overall co-ordination of the team with respect to timely meetings, reports, etc.; development of the research article, etc.
Analyst	1	Responsible for gathering all the requirements, tools, IDEs, materials, etc., necessary to design and implement the AI-based application; MUST work with the *Project Manager* and *Programmers* in developing the research article and developing the AI-based solution, respectively.
Programmer	2	Responsible for processing all the requirements and designs; and in turn, transform these inputs into codes that will yield the AI-based application; MUST work with the *Quality Control* personnel in developing appropriate test cases and/or scenarios.
Quality Control	1	Responsible for developing test cases with respect to unit/-component testing, integration testing, and acceptance testing; ensures that the AI-based application meets ALL the requirements stipulated herein.

8. **Evaluation Criteria:**

Table 2: Evaluation Criteria/Rubric with reference to AI-based Project

Criteria	**Weight** (%)
Data Preprocessing	15
Feature Engineering or Feature Learning	15
Data Transformation	10
Code Quality and Choice of ML/DL Algorithm	10
Performance/Evaluation Metrics	15
Research/Project Article	20
Presentation and Team Cohesion	15

9. **Rules for Acceptable Submission:**

(1) Each submission MUST be submitted as a compressed directory/folder (*.zip*) file via Canvas^{TM} LMS.

(2) Teams MUST use and maintain a consistent file-naming standard with respect to the constituent files of their AI-based Project directory/folder. Each team MUST rename their project folder to: "`project_{Group Name}`" prior to upload and submission via Canvas^{TM} LMS.

(3) Avoid spelling and/or grammatical errors

(4) Any form of advertisement whatsoever is unacceptable.

(5) No commercial/public URL(s), webpage(s), or website(s) of any kind will be accepted. A score of zero (0) will be awarded if this rule is broken.

(6) This is team project and task. Thus, each team's submission MUST be unique such that the contents of the project submission and the application resources MUST reflect the team's original work. Plagiarism of any form or kind will result in a score of zero (0).

(7) All the aforementioned rules are associated with *penalty points* when and/or where broken. Thus, those *penalty points* which have not been explicitly stated herein are subject to the instructor's discretion; and they will be reported to any defaulting team or student after grading, if/where necessary.

(8) Kindly review the Course Syllabus for the university's policies regarding Academic Integrity and Student Code of Conduct - https://sacramentostate.policystat.com/policy/11300038/latest/