

---


---

**SP-24-RED-DATA MINING**

**DISEASE PREDICTION SYSTEM: CANCER AND DIABETES**

**CS 4850 - Section 02 – Spring 2024**

**March 2024**

	 Kokou

**Team Members:**

<b>Name</b>	<b>Role</b>	<b>Cell Number</b>	<b>Email</b>
Alexus Glass	Developer	4049601998	lushgaloreinc2@gmail.com
Samuel Futral	Documentation	7067162064	samuelfutral@gmail.com
Kokou Adje	Developer	4703583315	adjelaime@gmail.com
Keegan Begley	Lead	6783667862	kbegley@students.kennesaw.edu
Sharon Perry	Project Advisor	7703293895	sperry46@kennesaw.edu

## 1. Overview:

“Disease Prediction System: Cancer and Diabetes” is a Senior class project which is focused on the prediction of two serious health conditions: cancer and diabetes. These two diseases have a huge impact on community health and represent one of the highest rates of mortality. Many studies are in progress to eradicate them, and early detection improves the effective treatment of the patients.

First, the objective of this project is to design machine learning models capable of predicting cancer and diabetes using patient’s health indicator information provided such as age, blood pressure, breast density, glucose level, etc. Secondly, develop easy user-friendly interface web application to gather input data from the user, and display the result in a real time manner. Finally, implement a strong data mining and processing to guarantee data and prediction accuracy.

The scope of this project consists of providing some deliveries as project plan, Gantt chart, requirements, analysis and design, reports, and requirements, analysis, and design.

## 2. Platform:

For this project, we will use a hardware and software platform. The right hardware needs to be chosen to properly run the machine learning program and must impact the quality and the performance of the models. The hardware requirements are the following for a computer to operate this application:

- ✓ Adequate processor for the execution of instructions.
- ✓ Good storage and memory to store a large amount of data from the processor.
- ✓ Graphics Processing Units (GPUs) to handle graphical data.
- ✓ For future utilization, high-performance computing resources like cloud infrastructure can be used for the software.

For the software infrastructure, the followings are required:

- ✓ Python: is important for the implementation of this project.
- ✓ Large Language Model (LLM): a kind of AI trained to recognize and generate text.
- ✓ Web frameworks: Flask, Django which offer some features for the application.

## 3. Collaboration Tools:

Communication	—	Cellphones (Call/Text) / Microsoft Teams / GroupMe
Collaboration	—	Discord (Mandatory unless another tool is authorized by Perry)
Version Control	—	GitHub

#### **4. Deliverables:**

- a. Team/Project Selection document (Individual Assignment)
- b. Weekly Activity Reports (WARs – Individual Assignment)
- c. Peer Reviews (Individual Assignment)
- d. Project Plan (Group Assignment)
- e. Project Requirements And Design
- f. Present Prototype for Peer Review (Group Assignment)
- g. Website (Group Assignment)
- h. Video Demo (Group Assignment)
- i. C-Day Application/Submission (Group applies to C-Day but each member submits individual bonus points documentation in Individual Assignments)
- j. Final Report Package (Group Assignment)
- k. iOS and Android compatible mobile time travel apps

#### **5. Project Schedule and Task Planning (GANTT CHART)**

02/21/2024

- Research dataset
- Create graphs
- Data analysis

02/28/2024

- Data mining
- Organize Data Table
- Data analysis
- Start on System Design

03/06/2024

- System Design & architecture
- Research
- Python Coding & SQL

03/13/2024

- System Design & Architecture
- Python Coding & SQL

03/20/2024

- Documentation & Report
- Python Coding & System structure

03/27/2024

- Documentation Report

04/03/2024

- Documentation & Code implementation

04/10/2024

- Finish Final report draft

04/17/2024

- Make sure the project is precise, clean and easy to understand.

04/21/2024

- Turn in the project.

## 6. Meeting Schedule

The team will be meeting on Mondays and Wednesdays after class at 8pm. The length of the meetings will be an hour typically, with the option of additional time if needed. During these meetings we will discuss our current focus, allocate tasks, and update project schedules based on progress status.

## 7. Statement of Participation

Project ID: SP-24

### STATEMENT OF PARTICIPATION:

By signing below, I Kokou Adje acknowledge that I will participate in all meetings, communications, deliverables, and other tasks necessary to complete the project. If I do not, I understand that Professor Perry will meet with me to remedy the situation.

Kokou Adje

3/7/2024

---

---

Team Member

Date

By signing below, I Hailey Walker acknowledge that I will participate in all meetings, communications, deliverables, and other tasks necessary to complete the project. If I do not, I understand that Professor Perry will meet with me to remedy the situation.

Alexus Glass

3/7/2024

---

---

Team Member

Date

By signing below, I William Stigall acknowledge that I will participate in all meetings, communications, deliverables and other tasks necessary to complete the project. If I do not, I understand that Professor Perry will meet with me to remedy the situation.

Samuel Futral

3/7/2024

\_\_\_\_\_

\_\_\_\_\_

Team Member

Date

By signing below, I William Stigall acknowledge that I will participate in all meetings, communications, deliverables, and other tasks necessary to complete the project. If I do not, I understand that Professor Perry will meet with me to remedy the situation.

Keegan Begley

3/7/2024

\_\_\_\_\_

\_\_\_\_\_

Team Member

Date

8. Gantt Chart

**Project Name:** SP-2 Red - Disease Prediction System: Cancer and Diabetes

**Report Date:** 02/18/2024

Deliverable	Tasks	Complete %	Current Status Memo	Assigned To	Milestone #1				Milestone #2				Milestone #3		C-Day	
					02/21	02/28	03/06	03/13	03/20	03/27	04/03	04/10	04/17	04/24	05/01	05/08
Requirements	Research dataset	0%	In progress	Alexus	10											
	Create graphs	0%		Kokou	10											
				Keegan	10											
	Data analysis	0%		Samuel	10											
Project design	Data mining	0%		Alexus		5										
	Organize Data Table	0%		Kokou		10										
	Data analysis	0%		Samuel&Keegan		5										
	Start on System Design	0%		Alexus		10										
Development	System Design & architecture	0%		Alexus & Kokou			10									
	Research	0%		Samuel&Keegan			10									
	Python Coding & SQL	0%		Alexus & Kokou			40									
Final report	Documentation & Report	0%		Samuel					8							
	Code implementation	0%		Alexus & Kokou					8	8	8					
	Finish Final report draft	0%		Samuel&Keegan								10				
	Project's review	0%		Alexus&Samuel &Kokou&Keegan								10	5			
	Turn in the project	0%		Alexus										5		
Total work hours					40	30	60	0	16	8	8	20	5	5	0	0

\* formally define how you will develop this project including source code management

Legend

Planned 192  
 Delayed  
 Number Work: man hours

