Synthesis Paper

For the Synthesis Prototype Project Demo, I created a software application for predicting diabetes. In developing this website, I required the knowledge and skillsets I gained from each curriculum area: Software Systems (SS), Business Analytics (BA), Data Management (DM), Cybersecurity and Networking (CN), especially from the first three areas.

Project Goal: The project goal of my synthetic project is to develop a user-friendly and effective website that enables medical professionals to predict the risk of diabetes in their patients accurately. This platform aims to empower doctors by providing them with advanced predictive analysis tools, allowing for early detection and intervention to improve patient outcomes and promote better health management practices.

Software Systems: For me, Software Systems encompasses both the theoretical principles and practical skills essential for designing, developing, and managing software applications. I was able to learn object-oriented paradigm and structured systems design to gain proficiency in procedural and object-oriented programming, database management, and web development. Now I have the knowledge and capabilities to design and implement software solutions effectively.

In this synthesis project, I started designing the software by outlining the software application's requirements, a process I learned from the Object-Oriented Analysis and Design course in the Software Systems curriculum. Based on the requirements I drew class diagram and use case diagram (UML diagrams) using draw.io. I used python programming language for the required coding and html CSS for the design of the interface of the website, which I learned from the Tech Foundation course. These knowledge and skills reflect the knowledge from the coursework Object Orient Analysis and Design from Software Systems (SS) curriculum.
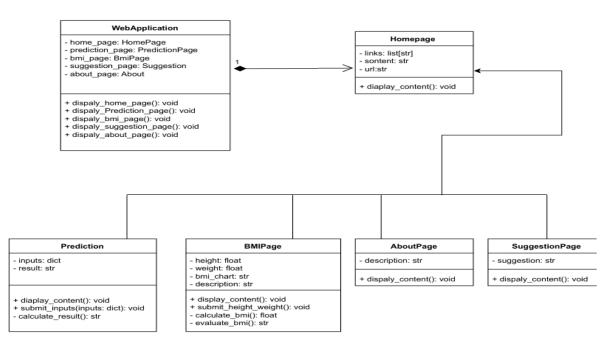
WebApplication

- home_page: HomePage
- prediction_page: PredictionPage
- bmi_page: BmiPage
- suggestion_page: Suggestion
- about_page: About

+ dispaly_home_page(): void
+ dispaly_Prediction_page(): void
+ dispaly_bmi_page(): void
+ dispaly_suggestion_page(): void
+ dispaly_about_page(): void

1

Homepage

- links: list[str]
- sontent: str
- url:str

+ diaplay_content(): void

Prediction

- inputs: dict
- result: str

+ diaplay_content(): void
+ submit_inputs(inputs: dict): void
- calculate_result(): str

BMIPage

- height: float
- weight: float
- bmi_chart: str
- description: str

+ display_content(): void
+ submit_height_weight(): void
- calculate_bmi(): float
- evaluate_bmi(): str

AboutPage

- description: str

+ dispaly_content(): void

SuggestionPage

- suggestion: str

+ dispaly_content(): void

Fig: Class Diagram

Use Case Diagram

Diabetes prediction application

Open software

Navigate to pages

Provide the inputs

Calculate diabetes prediction

Click on suggestion

Check BMI

Click on diabetes prediction

Learn about diabetes

include

Calculate the prediction result

Display prediction result

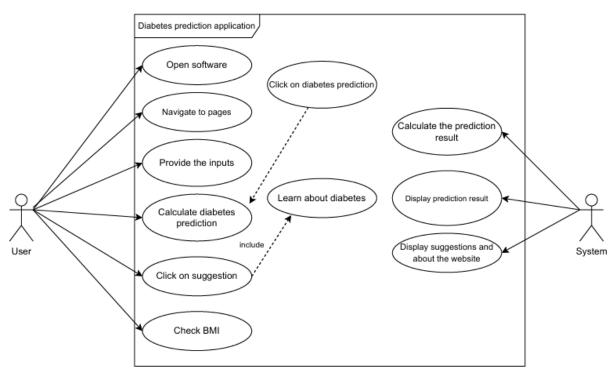Display suggestions and about the website

User

System

Fig: Use case Diagram

Data Management: Data Management (DM), to me, involves the extraction, transformation, and loading of data from various sources to use in data persistence platforms, as well as the study and application of data mining methods for business intelligence operations.

After drafting the project outline and website design, I started thinking about how I would approach the prediction. I learned about various machine learning models from the data management curriculum. Then, I began searching for an appropriate dataset for my model. I tried several datasets; some had too many variable columns with insignificant correlations to the outcome variable, others had many missing values, and some contained noisy data, which created anomalies in the results. From the Business Data ETL course, I learned how to handle missing data and remove irrelevant attributes. For this project, I had the option to choose the dataset myself, so I decided not to stress too much over data preparation and to focus more on building the website and applying machine learning. I wanted to select a more suitable dataset. I selected three datasets: one from the "Deep Learning Approach for Diabetes Prediction Using PIMA Indian Dataset," a diabetes prediction dataset from Kaggle.com: A Comprehensive Dataset for Predicting Diabetes with Medical & Demographic Data, and another from a class project. The dataset from the class project was already cleaned; using it would not allow me to demonstrate my data processing skills. So, I decided to use the second one. The link of my dataset is [Diabetes prediction dataset (kaggle.com)](kaggle.com) .

After selecting the dataset, I began the data cleaning and processing methods to prepare it for the machine learning model. This dataset contained nine variable columns, among which two—"gender" and "smoking_history"—were categorical rather than numerical. Since supervised ML can only handle numerical data and the "smoking_history" variable was highly correlated with the outcome, I converted these categorical attributes to numerical ones by mapping them with numbers. This is how I processed my dataset. From loading the dataset from the original source, importing necessary libraries, searching for missing values, performing correlation matrix analysis, cleaning, and processing the data, and finally saving the processed data into another CSV file—these tasks all integrate the skills from the Business Data ETL course of the DM era. For predicting the results, I used machine learning techniques: splitting the dataset into training and prediction datasets, training them, building the model, and finally using a random forest classifier for prediction. This encompasses the knowledge from the Data Mining course of this curriculum.

Business Analytics (BA) involves using statistical tools like logistic and linear regression, along with machine learning and predictive modeling, to derive insights from data for business decision-making. It also encompasses understanding information visualization techniques to present data effectively for analysis and decision-making.

In my project, I have several examples of concepts from the curriculum areas of Business Analytics. For instance, in a project focused on predictive modeling for diabetes prediction, I utilized machine learning tools such as random forest models to train data and make predictions, also model's accuracy analysis. Additionally, I applied statistical techniques like correlation matrix analysis and calculation of correlation coefficients to understand relationships between variables. Furthermore, for data visualization, I created bar charts to visually represent key insights derived from the data. This part of this project demonstrates how I integrated concepts from the curriculum areas of Business Analytics into practical applications, showcasing my proficiency in utilizing statistical tools and visualization techniques to derive meaningful insights from data.

Cybersecurity and Networking: Cybersecurity and Networking (CN) involves protecting computer systems and networks from cyber threats. This includes learning about different attack methods, security protocols, and how to detect intrusions. We also study topics like web security, cryptography, and the legal and ethical aspects of cybersecurity. In addition, we explore Digital Forensics and Fraud Detection, using techniques such as artificial intelligence and advanced machine learning to identify and prevent fraud in various domains like finance and healthcare.

The project must be an artifact the proves the synthesis of at least two curricular areas from the MS-CISBA program. My project already is the synthesis of three curricular areas. But my biggest support has come from this section. Not directly from securing the networking area, but from the technical part of it. In the Digital Forensics and Fraud course I created a website about identifying whether any organization's financial is fraudulent or not. Keeping in mind this concept, I created this Diabetes prediction website, but in an elaborative and more complex format.

In conclusion,

I used specific examples from this section to illustrate and support the development of my skills and knowledge throughout my academic and professional journey., which illustrates the application and integration of the curriculum areas through detailed project examples. For instance, the 'Diabetes Prediction Website' showcases my proficiency in Software Systems, Business Analytics, and Data Management by demonstrating my ability to design, develop, and effectively implement software solutions—highlighting the evolution of my skills from theoretical learning to practical application. This project, along with others documented in the 'Evolutions' material, not only reflects my academic accomplishments

but also outlines my professional growth trajectory, showcasing how each project undertaken has prepared me for complex real-world challenges and underscoring my readiness to contribute meaningfully to the field of computational information systems.