Ideation Phase

Defining the Problem Statements

Date	27-09-2023
Team ID	679
Project Name	Data Warehousing with IBM Cloud Db2 Warehouse

Data Warehousing with IBM Cloud Db2 Warehouse

Problem Definition and Design Thinking Introduction:

Heart disease is a prevalent and potentially life-threatening medical condition that affects millions of people worldwide. Early detection and risk assessment are crucial for preventing and managing heart diseases effectively. The task at hand is to data mine a Heart Disease data warehouse with Db2 warehouse in IBM cloud. To collect some of data set based on Heart Disease from different sources and combining into one data warehouse using ETL tool.

In this document, we will outline the problem statement, the steps involved in solving it, and the design thinking approach that will guide our project.

Problem Statement:

Objective: To bring together data from various sources, perform advanced data integration and transformation, and provide data architects with the tools to explore, analyse, and deliver actionable data for informed decision-making.

Data: We have multiple datasets containing various features for Heart Disease prediction (e.g., age, blood pressure, cholesterol levels, etc). These data will be used to integrate and transform into a robust data warehouse.

Key Challenges:

1. Data Warehouse Structure: Define the schema and structure of the data warehouse to accommodate various data sources.

2. Data Integration: Identify data sources and design a strategy to integrate data seamlessly into the data warehouse.
3. ETL Processes: Plan and implement ETL processes to extract, transform, and load data into the warehouse.
4. Data Exploration: Design queries and analysis techniques to empower data architects to explore and analyse data.
5. Actionable Insights: Focus on delivering actionable insights by enabling informed decision-making based on data.

Design Thinking Approach

1. **Empathize:**
   - Identify the stakeholders: Patients, doctors, hospitals, and data scientists.
   - Conduct interviews and surveys to understand their pain points and needs in heart disease prediction.
   - Gather insights about the data sources available for heart disease prediction, data quality, and existing challenges.

2. **Define:**
   - Clearly define the problem you aim to solve, such as improving the accuracy of heart disease prediction.
   - Create a user-centered problem statement based on your findings from the empathy phase.

3. **Ideate:**
   - Brainstorm potential data sources and variables that can contribute to heart disease prediction.
   - Explore different data warehousing solutions and architectures suitable for healthcare data.
   - Ideate on the features and functionalities that would benefit both healthcare providers and patients.
   - Prioritize ideas based on their potential impact and feasibility.

4. **Prototype:**

- Develop a data warehousing system that can efficiently store and manage healthcare data, including electronic health records (EHRs), medical images, and patient histories.
- Implement a predictive model using machine learning algorithms. Features might include patient demographics, lifestyle factors, medical history, and genetic data.
- Ensure data privacy and compliance with healthcare regulations like HIPAA.

5. **Test**:
    - Test the predictive model's performance using historical data. Evaluate its accuracy, sensitivity, specificity, and other relevant metrics.
    - Collect feedback from healthcare providers and patients on the usability and effectiveness of the system.
    - Refine the model based on feedback and iterate on the data warehousing solution for scalability and efficiency

6. **Implement**:
    - Deploy the finalized predictive model and data warehousing system in a healthcare setting, ensuring seamless integration with existing infrastructure.
    - Train healthcare providers and staff on using the system effectively.
    - Monitor the system's performance in real-world scenarios and make necessary adjustments

7. **Iterate**:
    - Continuously collect and analyze new data to improve the predictive model's accuracy and effectiveness.
    - Gather user feedback and make iterative improvements to the data warehousing system and user interface.
    - Stay up-to-date with advances in healthcare and machine learning to incorporate new features and technologies.

8. **Scale and Evolve**:
    - As the system proves its value, consider scaling it to other healthcare facilities or regions.

- Explore partnerships with other healthcare organizations and researchers to expand the dataset and improve predictions.
- Keep the system adaptable to changing healthcare needs and regulations.

Conclusion

In this document, we've outlined our approach to solving the problem of data mining a Heart Disease prediction model with db2 warehouse. We've defined the problem, identified key challenges, and laid out a design thinking approach that involves empathizing with users, defining objectives, ideating potential solutions, prototyping, testing, implementing, and iterating.

By following this approach, you can ensure that the Heart Disease data warehouse meets the needs of its users and delivers valuable insights for Cardiology research and patient care while using Db2 Warehouse on IBM Cloud as the underlying technology infrastructure.