Disease Progression Challenge

Goals

- Overview of your design including at least 1 slide with the architecture described.
- Talk through what assumptions and considerations you used to arrive at this design.
- Pros and cons of the proposed design.
- Be prepared to answer questions from each interviewer.

Description

Your company specializes in developing software solutions for healthcare providers to improve patient care and outcomes for those diagnosed with lung cancer or COPD. One of your clients has asked for a new system that can help them identify patients at risk of disease progression, so that they can intervene early and prevent worsening of symptoms or the development of complications. However, the client's patient data is spread across multiple systems and is stored in a variety of formats, including electronic health records (EHRs), laboratory information systems (LIS), claims data, and imaging studies.

Your challenge is to design a data architecture that can support this goal. Specifically, you need to:

1. Identify the disparate data sources that will be needed to support this initiative and design a data integration strategy that can extract and transform the relevant data into a standardized format. This will include developing a strategy for handling imaging studies, such as CT scans, MRIs, or PET scans, which may require specialized tools and techniques.

2. Design a data model that can store and organize the integrated patient data in a way that supports the necessary queries and reporting, while also ensuring data quality, privacy, and security.

3. Develop a set of data pipelines that can collect, process, and transform the necessary data from each source, while ensuring that the data is up-to-date and consistent. This will include developing a strategy for handling imaging studies, which may involve using algorithms to extract relevant features from the images.

4. Articulate (at a high level) what machine learning algorithms might be used that can analyze the integrated patient data to identify those at high risk of disease progression, while considering the unique features and challenges of each data source, including imaging studies.

5. Mockup a system that provides a user-friendly dashboard that healthcare providers can use to view patient risk scores, and provide actionable insights to improve patient care and outcomes.

You will need to consider a range of complex factors in designing this data architecture, including data integration challenges, data quality issues, scalability, performance, regulatory compliance, and user experience.