Hello, my name is Shuo Wang, and I'm excited to tell you about a complex AIGC application that I worked on and the role I played in its development.

I was part of a team that was tasked with developing a system to automatically generate summaries from medical notes in XML format and introduce several passive digital markers (PDMs) that can help reduce the burden of doctors  
by automating the assessment using medical notes. Our goal was to extract the most important information from these notes and present it in a concise, easy-to-understand format.

To achieve this, we first had to engineer a data transformation pipeline that could convert the complex XML data into a more manageable format. This involved working with medical experts to understand the relevant data points and developing algorithms to extract them from the XML.

Once we had the data in a more structured format, we utilized DF-ICF topic modeling to extract the top 50 topic words from each note. This allowed us to identify the most important concepts discussed in each note and summarize them in a concise text summary.

We then implemented three cutting-edge DNN models (Ensemble BERT, Sequence BERT, Word BERT) to effectively categorize remission of schizophrenia and bipolar disorders into three categories - severe impairment, serious impairment, and moderate to no impairment. This was an important part of the application because it allowed us to categorize patients based on the severity of their condition and provide tailored treatments.

Throughout the development process, we faced several challenges.

One of the major challenges we encountered was dealing with the complexity of the medical notes. The XML format made it difficult to extract the relevant information needed for the data transformation pipeline. We also had to deal with inconsistencies in the data, such as typos, misspellings, and abbreviations. To address these issues, we had to develop a series of data cleaning and normalization techniques to ensure the data was accurate and consistent.

Another challenge we faced was optimizing the DNN models to effectively categorize the remission of schizophrenia and bipolar disorders. We had to experiment with different models and parameter settings to achieve the highest accuracy and F1-score possible. Additionally, we had to ensure that the models were not overfitting to the training data, which required extensive validation and testing.

Lastly, we had to ensure that the results of our models were interpretable and actionable for the healthcare professionals who would be using them. This required collaboration with a cross-functional team to analyze and interpret the complex model results and translate them into actionable insights. We also had to ensure that the model results were presented in a clear and concise manner that could be easily understood by non-technical stakeholders.

Overall, I played a key role in developing this AI-Generated Content application. My contributions included developing the data transformation pipeline, implementing the DNN models, and collaborating with the team to analyze and interpret the results. Despite the challenges we faced, we were able to develop a robust and effective system that has the potential to significantly enhance the diagnosis and treatment of mental health disorders.

Thank you for your time.