**Project**: Investigation of Transformer models for implementation in NFL Scouting Report Generator

**Team Members and Contributions**:

1. Jason Jordan: Implementation of vision transformer for assessment of image content
2. Krutarth Lad: Attention visualization with BERT
3. Harshitha Behera: Enhanced Chatbot with DialoGPT
4. Jainil Patel: Question Answering system refinement

**AI Techniques Utilized with Results and Challenges for each**:

The purpose of this hands-on project is to evaluate various transformer models to determine their suitability for use in the CSC 5542 class project, which seeks to analyze football player performance and mobility data to generate a scouting report to assist teams with game preparation and strategy modification. The types evaluated include a vision transformer, an attention visualization using BERT, and an enhanced chatbot with DialogGPT.

*Vision Transformer*:

The vision transformer (ViT) was utilized to analyze football-related images, such as live-action video, game play still shots, stadium scenery, and/or play formation diagrams. The ViT used in this project was the ‘Vit Large Patch16 384’ which was pre-trained on a dataset of over 14 million images for classification into one of the 1,000 ImageNet classes. A video link is provided that shows the operation of the program, which creates a Streamlit application to classify a given image and allows a user to upload an additional image for classification.

The application was somewhat successful in that the first image was classified as a “football helmet” which is partially correct given that players in the scene were wearing helmets. However, this may not be suitable for the final research project as a more thorough scene description is desired instead of predicting a single class from the ImageNet classes, many of which are unrelated to American football. An additional image of play formation diagram was also uploaded in the video and the result was even worse, with the model classifying the image as an “analog clock.” This project demonstrates the ability to deploy a ViT on a user-end application, but a more powerful model, perhaps with additional fine-tuning, will be necessary for better scene description.

*Attention Visualization with BERT*:

Attention visualization with BERT can significantly enhance the analysis of language in NFL game strategies and player performance. By understanding how BERT's attention mechanism focuses on specific words and phrases, we can gain insights into the underlying patterns and contexts in play descriptions, player interviews, and social media sentiment.

In the context of the NFL Super Bowl project, this technique allows us to identify key terms related to game strategies, player strengths, and weaknesses, helping teams make informed decisions. For example, when analyzing play-by-play commentary, we can visualize which words BERT prioritizes, revealing critical factors influencing game outcomes. This could include the importance of specific player names, actions, or even opposing team strategies.

Additionally, attention visualization can aid in sentiment analysis by pinpointing phrases that evoke strong emotional responses from fans or analysts. By correlating these insights with real-time data from the game, we can create a dynamic and adaptive strategy report that highlights essential focus areas for coaches and players.

Ultimately, using attention visualization in conjunction with BERT enhances our understanding of complex linguistic patterns in NFL contexts, facilitating more effective decision-making and strategic planning.

*Enhanced Chatbot with DialogGPT*:

An enhanced chatbot allows detailed interactions which users can process in real time. This is critical for strategy development and adjustments during sporting events. Teams can request situational input from the chatbot and can pose additional questions to refine the output into a practical application, all while preserving the overall history of the interaction.

The chatbot developed for this project demonstrates reasonable performance for basic conversational requests. It’s suitability for domains related to American football is, however, not appropriate for immediate use in providing game strategy. A more refined model will need to be utilized that has been trained in a related domain, or the current model will need additional fine-tuning to work for our purpose.

**Suggestions for Further Improvement**:

The vision and chatbot models both have strong potential for use within the confines of the NFL Scouting Report Generation model since it will be incorporating both text and images/video. We believe that refinement of the model for our sporting domain will be key in their implementation. Therefore, further enhancements, such as fine-tuning with the appropriate information, will be necessary. The attention visualization aspect will be helpful in the refinement process.

Vision Transformer Demonstration:

<https://umsystem.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=5bd5898a-6a13-4531-b11a-b28b00129987>

BERT Attention Visualization Demonstration:

<https://umsystem.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=193be32f-ed23-4be5-a515-b28b0159ef8b>