**DSAN 6600 Final Project Proposal**

*Group: Hongxin Wu, Yanyan Li, Xinran Zhang*

1. **Project Summary**

Our project aims to address the limitations of current autocorrect systems by developing a neural network-based model that can detect and correct homophone errors within text. Our goal is to analyze the context surrounding homophones and employ advanced language models to discern and rectify incorrect usage.

* We have noticed that many auto-correct tools, such as Apple's in text messages, lack the ability to correct when an incorrect homophone is used. For example, all of the words in "We one the game!" are spelled correctly, but the wrong "won" was used.
* Some APPS use the wrong their/there/they're, petal/pedal, son/sun, knight/night, etc.

The context around these words should make which spelling is correct apparent, enabling context-based language models to predict the correct spelling. We envision this tool to process significant text inputs, identify homophone errors, suggest corrections, or even provide the corrected version directly. This tool, hosted on a Shiny dashboard or similar platform, will offer assistance and potentially integrate predictive text suggestions. This project is exciting due to its practical application in everyday communication, where such errors are common yet often overlooked due to cognitive load or oversight.

1. **Modeling Approach**

* We will utilize a pre-trained causal language model from Hugging Face, focusing on generative models for their prowess in predictions. These models excel in contextual understanding, which is crucial for homophone correction.

1. **Data Selection**

* Our primary resource will be a pre-trained language model from HuggingFace, selected for its size (≤100 million parameters) and generative nature.
* We might have to filter out non-printable characters, normalize text, and tokenize into words during preprocessing.

1. **Model Evaluation**

* Diverse Accent Testing: Evaluate the model with a diverse corpus of recordings, emphasizing inclusivity for various global accents to ensure wide applicability.
* Grammatical Error Analysis: Selection of test recordings will be meticulous, prioritizing those with common grammatical pitfalls to assess model correction capabilities.
* Performance Metrics: Assessing the model using accuracy, precision, recall, speed, coverage, Word Error Rate (WER), Error Detection Rate, and Error Correction Rate.

1. **Others**

* Web Application: Development of a user-friendly web interface for model interaction, potentially via Streamlit.
* Optimal Model Selection: The quest for the most efficient pre-trained language models will be pivotal. We'll explore the Transformers and Hugging Face ecosystems to find models that promise a compact footprint without sacrificing performance or accuracy.

1. **Project Challenges**

* Choosing a suitably robust yet compact pre-trained model is critical for the project's success.
* Implementing Beam Search for grammatical correction poses a significant challenge and will require innovative coding and algorithm development.
* Fine-tuning the model on a specialized dataset for homophone correction will be computationally extensive and require a lot of computating resources.
* Rule-based systems may not adapt well to the nuances of natural language, leading to rigid corrections that don't fit all contexts.

1. **Conclusion**

The successful implementation of this project has the potential to significantly enhance the accuracy and reliability of voice to text transformation and autocorrect systems. By leveraging deep learning techniques and advanced models, we aim to create a sophisticated tool that not only transforms the voice recording into text but also enhances the overall quality of the texts generated.