# Adversarial Attacks and Defenses in Speech Recognition Systems

## Abstract:

Adversarial attacks pose significant challenges to the robustness of speech recognition systems, potentially leading to misinterpretation or failure in real-world applications. This project aims to explore the vulnerabilities in current speech recognition models by analyzing different adversarial techniques targeting these systems. The focus will be on identifying the types of attacks that compromise speech recognition accuracy, such as targeted noise and perturbations. Additionally, the study will investigate defense mechanisms, including adversarial training and model modifications, to enhance the resilience of speech recognition models against such attacks. The outcome will contribute to developing more secure and reliable speech recognition technologies.

## Problem Statement:

Speech recognition systems are increasingly integrated into applications such as virtual assistants, healthcare, and security systems. However, these systems are vulnerable to adversarial attacks, where small, imperceptible changes to audio inputs can lead to incorrect transcriptions or command misinterpretation. This poses a significant risk to the reliability and security of speech-based applications. Current models fail to generalize effectively in the presence of these attacks, leading to a pressing need for robust defense strategies. The project seeks to study the various adversarial methods employed to exploit these vulnerabilities, assess their impact on speech recognition accuracy, and propose effective countermeasures to safeguard the integrity of these systems.