Our project aims to improve the performance of Arabic dialect identification by fine-tuning the wav2vec2 XLSR model. To enhance the model's capabilities, a small DNN layer was added on top of the wav2vec2 model. This resulted in the creation of a semi-streaming model, which clusters dialect chunks with a time mismatch of less than 1 second. The model was tested on a dataset of 5 different Arabic dialects, achieving an accuracy of 65%. This performance was higher than previous studies that used deep CNN models and different feature extraction methods, such as i-vectors. The success of this project highlights the potential of fine-tuning the wav2vec2 XLSR model for Arabic dialect identification and provides a new approach to the task. The results of this project provide a novel approach to the task of Arabic dialect identification and demonstrate the effectiveness of fine-tuning the wav2vec2 XLSR model for this task.