

## **Abstract**

Spoken language identification (LID) is the process of identifying the language spoken within an utterance. This task should be performed without knowing any information about the speech utterance like its contents or its speaker. Spoken LID has numerous applications, one important application arises in call centers dealing with speakers speaking different languages. It would be very applicable to be able to identify the language of a speaker within only a few seconds of his speech, then direct him to the corresponding recognition system or call center employee to answer his query. Another application is indexing large speech data archives that contain multiple languages. Accurate LID here leads to more faster and better search results. In this research, We use Support Vector Machines (SVMs), Ergodic Hidden Markov Models (HMMs), and Deep-Networks to achieve the task of spoken LID. For SVM approach, we extract acoustic features from speech utterances then, we build a two-class SVM classifier for each pair of languages. This approach is based on modeling acoustic properties of each language, but for HMM approach, we investigate different HMM systems for spoken LID. Ergodic HMM models are used so that each state models some distinct sounds in the language. In this approach, the state probabilities model the acoustics of the language while the state transition probabilities and to some extent the mixture weights model some phonetic properties of the same language. Doing this, we try to combine both the acoustic and the phonetic approaches which are used in the literature. While, in the Deep-Networks methodology, we reviewed the use of Deep-Networks as a classifier. We extend our work to differentiate between Arabic dialects like Egyptian Arabic, Iraqi Arabic, and Gulf Arabic.

We use Liblinear package which is a large scale and efficient software for SVM classification. For HMM experiments we use MATLAB to implement HMM training and testing.

## ملخص الرسالة

تحديد اللغة المنطوقة (LID) هو عملية التعرف على اللغة المحكية داخل الكلام. يجب أن يتم تنفيذ هذه المهمة من دون معرفة أي معلومات عن الخطأ. الكلام مثل محتوياته أو المقصود. و تحديد اللغة المنطوقة له العديد من التطبيقات، أحد التطبيقات المهمة ينشأ في مراكز الاتصال التعامل مع المتكلمين يتكلمون بلغات مختلفة. سيكون من العملي جدا القدرة على التعرف على اللغة من المتكلم في غضون ثوان قليلة فقط من كلامه ، ثم المباشرة به. وبين نظام التعرف على الاتصال او على مركز للموظف للاجابة على ما يريد الاستعلام عنه. تطبيق آخر هو فهرسة محفوظات البيانات الكبيرة التي تحتوي على كلمات متعددة اللغات. دالة تحديد اللغة هنا يؤدي الى مزيد من اسرع والفضل نتائج البحث. في هذا البحث استخدمت الآلات لسجلات دعم (SVMs) ونماذج ماركوف ار جوديك (Ergodic HMMs) والشبكات العصبية العميقة لتحقيق مهمة تحديد اللغة المنطوقة. لنهج SVM، كما تم استخراج الميزات الصوتية من الكلام المنطوق ، ثم تم بناء المصنف SVM المطبقين لكل زوج من اللغات. ويستند هذا النهج على نمذجة الخصائص الصوتية لكل لغة، ولكن بالنسبة لنهج نماذج ماركوف ار جوديك ، تقوم بالبحث في عدة نظم نماذج ماركوف ار جوديك مختلفة لتحديد اللغة المنطوقة . تستخدم نماذج ماركوف ار جوديك بحيث تعين كل حالة نماذج عن بعض الأصوات المتغيرة في اللغة. في هذا النهج ، احتمالات الحالات تقوم بنمذجة الصوتيات للغة، بينما تقوم احتمالات حالات الانتقال والتي حدد ما يخطط الأوزان بنمذجة بعض الخصائص اللفظية للغة نفسها. بفعل ذلك ، نحاول الجمع بين كلا من النهج الصوتية والنهج اللفظية التي تستخدم في البحث. في حين ، في نهج الشبكات العصبية العميقة تم امكانية استخدام هذه الشبكات بالميزات الصوتية لتعمل كمصنف. كما امتد عملنا لتفريق بين اللهجات العربية مثل العربية المصرية، العربية العربية ، والخليجية العربية.

تستخدم حزمة Liblinear وهو نطاق واسع وبرامج فعالة لتصنيف SVM. و نستخدم MATLAB لتقليد

تجربيات واختبارات نماذج ماركوف ار جوديك.