Low Resource Point Process Models for Keyword Spotting Using Unsupervised Online Learning

Samik Sadhu  Prasanta Kumar Ghosh
SPIRE LAB, Indian Institute of Science, Bangalore 560012, India

Problem Statement

Only as few as 100 languages in the world have annotated datasets for training Automatic Speech Recognition (ASR) and Keyword Spotting (KWS) systems. We are hence motivated to build a KWS system which can be trained with as minimum number of training samples as possible.

Point Process Model (PPM) based KWS

A new set of piece-wise constant rate-parameters is obtained

\[ \theta_k^0 = \frac{1}{M_k^0} \sum_{p \in P} \left| t_{k_{start}}^{p} - t_{k_{start}} \right| \]

\[ \theta_k^1 = \frac{1}{M_k^1} \sum_{p \in P} \left| t_{k_{start}}^{p} - t_{k_{start}} \right| \]

\[ \theta_k = (\theta_k^0, \theta_k^1) \]

We choose the value of \( a(k) \) to be

\[ a(k) = \frac{y(k) - y_{\text{median}}}{y_{\text{median}}} \]

Experimental Setup

We use 8 keywords from TIMIT database [3] with the TIMIT train corpus as the online learning corpus. The test corpus consists of all the sentences spoken by 24 speakers in the test TIMIT test corpus and the sentences spoken by 50 speakers in the development set defined by Kaliki [4] TIMIT recipe. \( P_{\text{ROC}} \) is calculated as area under ROC curve normalized by its support X100.

References


Conclusions and Future Work

The performance \( (P_{\text{ROC}}) \) of the unsupervised online learning algorithm tends towards that of the ordinary PPM algorithm as more number of sentences are observed from the online learning corpus.

Unsupervised online learning of PPM achieves a \( P_{\text{ROC}} \) performance 2% less than ordinary PPM with just 1% of the training data required for ordinary PPM.

Thus the algorithm is useful for situations where very less number of training samples are available.

Such class of algorithms need to be extended to other KWS algorithms and ASR systems.

The proposed algorithm should be verified on low resource language databases.