

Erinç Dikici
Thesis Supervisor: Assoc. Prof. Murat Saraçlar
SUPERVISED, SEMI-SUPERVISED AND UNSUPERVISED METHODS
IN DISCRIMINATIVE LANGUAGE MODELING
FOR AUTOMATIC SPEECH RECOGNITION

Abstract

Discriminative language modeling aims to reduce the error rates by rescoreing the output of an automatic speech recognition (ASR) system. Discriminative language model (DLM) training conventionally follows a supervised approach, using acoustic recordings together with their manual transcriptions (reference) as training examples, and the recognition performance is improved with increasing amount of such matched data. In this thesis we investigate the case where matched data for DLM training is limited or not available at all, and explore methods to improve ASR accuracy by incorporating unmatched acoustic and text data that come from separate sources. For semi-supervised training, we utilize weighted finite-state transducer and machine translation based confusion models to generate artificial hypotheses in addition to the real ASR hypotheses. For unsupervised training, we explore target output selection methods to replace the missing reference. We handle discriminative language modeling both as a structured prediction and a reranking problem and employ variants of the perceptron, MIRA and SVM algorithms adapted for both problems. We propose several hypothesis sampling approaches to decrease the complexity of algorithms and to increase the diversity of artificial hypotheses. We obtain significant improvements over baseline ASR accuracy even when there is no transcribed acoustic data available to train the DLM.

PUBLICATIONS

Journals

1. Dikici, E., Semerci, M., Saraçlar, M., Alpaydın, E., “Classification and Ranking Approaches to Discriminative Language Modeling for ASR”, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 21, No. 2, pp. 291-300, 2013. (SCI-E)
2. Dikici, E., Saraçlar, M., “Semi-supervised and unsupervised discriminative language model training for automatic speech recognition”, Elsevier Speech Communication, *accepted for publication*. (SCI)

Conferences - International

1. Dikici, E., Semerci, M., Saraçlar, M., Alpaydın, E., “Data Sampling and Dimensionality Reduction Approaches for Reranking ASR Outputs Using Discriminative Language Models”, 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), pp. 1461-1464, Florence, Italy, 28-31 August 2011.
2. Çelebi, A., Sak, H., Dikici, E., Saraçlar, M., Lehr, M., Prud’hommeaux, E., Xu, P., Glenn, N., Karakos, D., Khudanpur, S., Roark, B., Sagae, K., Shafran, I., Bikel, D., Callison-Burch, C., Cao, Y., Hall, K., Hasler, E., Koehn, P., Lopez, A., Post, M., Riley, D., “Semi-Supervised Discriminative Language Modeling for Turkish ASR”, 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), pp. 5025-5028, Kyoto, Japan, 25-30 March 2012.
3. Dikici, E., Çelebi, A., Saraçlar, M., “Performance Comparison of Training Algorithms for Semi-Supervised Discriminative Language Modeling”, 13th Annual Conference of the International Speech Communication Association (Interspeech 2012), pp. 206-209, Portland, OR, USA, 9-13 September 2012.
4. Dikici, E., Prud’hommeaux, E., Roark, B., Saraçlar, M., “Investigation of MT-based ASR Confusion Models for Semi-Supervised Discriminative Language Modeling”, 14th Annual

Conference of the International Speech Communication Association (Interspeech 2013), pp. 1218-1222, Lyon, France, 25-29 August 2013.

5. Dikici, E., Saraçlar, M., “Unsupervised Training Methods for Discriminative Language Modeling”, 15th Annual Conference of the International Speech Communication Association (Interspeech 2014), pp. 2857-2861, Singapore, 14-18 September 2014.
6. Dikici, E., Saraçlar, M., “MT-based Artificial Hypothesis Generation for Unsupervised Discriminative Language Modeling”, the 23rd European Signal Processing Conference (EUSIPCO 2015), pp. 1416-1420, Nice, France, 31 August - 4 September 2015.
7. Dikici, E., Saraçlar, M., E. Arısoy, “A Decade of Discriminative Language Modeling for Automatic Speech Recognition”, Lecture Notes in Artificial Intelligence, vol. 9319, pp. 11-22.

Conferences - National

1. Dikici, E., Saraçlar, M., “Curriculum Based Discriminative Language Model Training”, IEEE 21st Signal Processing and Communications Applications Conference (SIU 2013), Girne, North Cyprus, 2013.
2. Dikici, E., Saraçlar, M., “Unsupervised Discriminative Language Model Training”, IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014), pp. 1158-1161, Trabzon, Turkey, 2014.

Defense Jury Members

- | | |
|-------------------------------|---------------------|
| 1. Assoc. Prof Murat Saraçlar | Boğaziçi University |
| 2. Prof. Ethem Alpaydın | Boğaziçi University |
| 3. Prof. Levent Arslan | Boğaziçi University |
| 4. Assoc. Prof. Hakan Erdoğan | Sabancı University |
| 5. Assoc. Prof Engin Erzin | Koç University |

Defense Date: 21.06.2016