

RNN Language Model

- RNN-LM has lower *perplexity* and *word error rate* than 5-gram model with Knesser-Ney smoothing.
- The experiment is held on Wall Street Journal corpus:

Model	# words	PPL	WER
KN5 LM	200K	336	16.4
KN5 LM + RNN 90/2	200K	271	15.4
KN5 LM	1M	287	15.1
KN5 LM + RNN 90/2	1M	225	14.0
KN5 LM	6.4M	221	13.5
KN5 LM + RNN 250/5	6.4M	156	11.7

- Later experiments: char-level RNNs can be very effective!

Mikolov, Karafiát, Burget, Cernocký, and Khudanpur. Recurrent neural network based language model. INTERSPEECH 2010.