

obtienen buenos resultados en media frente a los tres corpus de test.

Y por último, la clara correlación (aunque no estrictamente directa) entre el porcentaje de palabras OOV y la WER obtenida evidencia la necesidad de mejorar o adaptar los vocabularios de nuestros modelos.

Como líneas futuras, en recientes estudios se comienza a utilizar las redes neuronales para el entrenamiento de modelos de lenguaje. Aunque se sigue utilizando el entrenamiento “clásico” basado en la predicción estadística como primer paso, se realiza un rescoring de los modelos usando redes neuronales. Los resultados de este rescoring muestran mejoras sobre los resultados experimentales.

Otras posibles líneas de investigación serían aumentar el orden de los n-gramas con los que se entrenan los modelos de lenguaje, puesto que la tecnología actual ya permite trabajar con tetragramas (4-gramas) o incluso pentagramas (5-gramas); o disponer de modelos basados en el habla para poder utilizarlos de forma opcional según el carácter del corpus a reconocer.

Agradecimientos

El trabajo realizado está enmarcado en el proyecto del Plan Nacional TraceThem TEC2015-65345-P y en la red gallega TecAnDaLi ED431D 2016/011 financiada por la Xunta de Galicia. Asimismo se beneficia de las ayudas de la Xunta de Galicia de Grupos de Referencia Competitiva GRC2014/024 y Agrupación Estratégica Consolidada de Galicia acreditación 2016-2019 y a la Unión Europea a través de los fondos FEDER. Se agradece al Instituto Ramón Piñeiro de la Xunta de Galicia el acuerdo de colaboración para la utilización del material del CORGA y su participación en el etiquetado de los corpus 2 y 3.

Referencias

Peddinti, Vijayaditya, D. Povey y S. Khudanpur. 2015. A time delay natural network architecture for efficient modeling of long temporal context. En *Proceedings of INTERSPEECH*.

Stolcke, Andreas. 2002. SRILM An extensible language modeling toolkit. En *Proceedings of the International Conference on Statistical Language Processing*. Denver, Colorado.

García, Carmen, J. Tirado, L. Docío y A. Cardenal. 2004. Transcrigal: A bilingual system for automatic indexing of broadcast

news. *IV International Conference on Language Resources and Evaluation*.

Docío, Laura, A. Cardenal y C. García. 2006. TC-STAR 2006 automatic speech recognition evaluation: The uvigo system. En *Proc. Of TC-STAR Workshop on Speech-to-Speech Translation*. ELRA, París, France.

Jurafsky, Daniel, y J.H. Martin. 2008. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*.

Vicente, Marta, C. Barros, F. Peregrino, F. Agulló y E. Lloret. 2015. La generación de lenguaje natural: análisis del estado actual. *Computación y Sistemas*. Volumen: 9, n.º 4.

Povey, Daniel, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlíček, Y. Quian, P. Schwarz, J. Silovský, G. Stemmer y K. Veselý. 2011. The Kaldi Speech Recognition Toolkit. En *ASRU Workshop*.

Campillo, Francisco y E. Rodríguez. 2005. Evaluación del modelado acústico y prosódico del sistema de conversión texto-voz Cotovía. En *Procesamiento del Lenguaje Natural*. Volumen 35, páginas 5-12.

Alegría, Iñaki, I. Arantzabal, M. Forcada, X. Gómez, L. Padró, J.R. Pichel y J. Waliño. 2006. OpenTrad: Traducción automática de código abierto para las lenguas del estado Español. En *Procesamiento del Lenguaje Natural*. Volumen: 37, páginas 356-358.

Mikolov, Tomas, S. Kombrink, A. Deoras, L. Bruget y J. Cernocky. 2011. Rnnlm-recurrent neuronal network language modeling toolkit. En *Proc. of ASRU Workshop*.

Xu, Hainan, T. Chen, D. Gao, Y. Wang, K. Li, N. Goel, Y. Carmiel, D. Povey y S. Khudanpur. 2018. A pruned rnnlm lattice-rescoring algorithm for automatic speech recognition. En *ICASSP*.

Sundermeyer, Martin, Z. Tüske, R. Schlüter y H. Ney. 2014. Lattice decoding and rescoring with long-span neural network language models. En *Fifteenth Annual Conference of the International Speech Communication Association*.

Chen, Xie, X. Liu, A. Ragni, Y. Wang y M. Gales. 2017. Future word contexts in neural network language models. ArXiv preprint arXiv:170805592.