Deep learning approach for autonomous medical diagnosis in spanish language

GJ. Daquarti[1], AE. Alfonso[1], F. Nanni[1], H. Ferrero[1], F. Murzone[1], AM. Groisman[1], F. Arias[1], J. Estevez[1]
[1] UMA, Buenos Aires, Argentina

Background.
Autonomous medical diagnosis may improve healthcare and save cost. Spanish language has over 483 million native speakers and it is the second used language in the United States. Available tools for artificial intelligence diagnosis in spanish are rule based systems with poor accuracy. The aim of this work was to develop an autonomous medical diagnosis tool for spanish speakers using a deep learning approach.

Model.
Anonymized data of 759000 general practice electronic health records were analyzed. A secuencial model was implemented (figure) :
A) Symptoms translator: The input was a text-sentence that represent a colloquial symptom from a patient, while the output was a medical term. B) Symptoms extractor: The input was A while the output was a list of medical symptoms. C) Text generator: The input was B while the output was a paragraph of two text-sentences originated from patient's answers. D) Diagnosis classifier: The input was C and the output was a list of 79 possible medical diagnosis. Hierarchical Attention Network model approach was implemented. The layers were: 1) Embedding, 2) Bidirectional recurrent neural networks using gated recurrent units 3) AttentionWithContext 4) Softmax. Top-1 diagnosis accuracy on validation set was 82% and 93% in Top-3. The model was develop with Python libraries in a Keras-Tensorflow framework.

Implementation.
A minimum viable product is available online on https://uma-health.com/autonomous

Conclusion.
Our autonomous medical diagnosis model in spanish language may provide a fully accessible tool to improve healthcare and optimize medical resources.