Grammatical gender in Swedish is predictable using recurrent neural networks

Edvin Listo Zec[1], Olof Mogren[1]

[1]*RISE Research Institutes of Sweden*

Presenter's e-mail address: edvin.listo.zec@ri.se, olof.mogren@ri.se

The grammatical gender of Swedish nouns is a mystery. While there are few rules that can indicate the gender with some certainty, it does in general not depend on either meaning or the structure of the word. In this work we demonstrate the surprising fact that grammatical gender for Swedish nouns can be predicted with high accuracy using a recurrent neural network (RNN) working on the raw character sequence of the word, without using any contextual information.

There are two different grammatical genders for Swedish nouns, *common (utrum)* and *neuter (neutrum) (Rebbe, Gullberg & Ivan, 1954)*. Swedish used to have more than two genders; however the masculine and the feminine later merged into the common gender and modern Swedish makes no difference between them. The grammatical gender affects the indefinite article (*en* boll -- *ett* bord; *English: a ball -- a table*), as well as the definite article and the definite suffix. The Institute for Language and Folklore, a Swedish government agency, states that there is no unequivocal way of determining the grammatical gender of a word. Further the Swedish Academy states that the grammatical gender in Swedish is a lexical property which is to a large extent not dependent on neither meaning nor the structure of the noun (Teleman & Hellberg, 1999). There are exceptions: living things are generally utrum (the more common class), and there are a few specific suffixes that are reasonably good predictors.

However, determining the gender of a word is generally considered a major difficulty for non-native Swedish speakers learning the language, and it is an open question how predictable it is.

In this paper, we investigate the predictability of grammatical gender for Swedish nouns without any context words. We train a recurrent neural network (RNN) model to predict the property given only the raw character sequence as input. The model quite easily learns the necessary patterns and achieves a 95% accuracy on the test set, which should be compared to the 71% majority class baseline.

In related work, (Nastase & Popescu, 2009) trained a kernel ridge regression (KRR) model to predict the gender of nouns in German and Romanian to a high precision. It is however well known that Romanian has phonological patterns that indicate gender, while German is more similar to Swedish. Previous work has been done for predicting the grammatical gender of Swedish nouns using neural networks (Basirat & Tang, 2018). However, here the authors used word embeddings with context information from a corpus and achieve a high classification accuracy of 93.7% on their test data. Our work differs from theirs since we do not depend on any such corpus; our models act solely on raw character sequences.

A recurrent neural network (RNN) is a class of artificial neural networks that can model a sequence of arbitrary length, using weight sharing between each position in the sequence. Unrolling the RNN results in a feed-forward neural network. In the basic RNN variant, the transition function at time step t is a linear transformation of the hidden state at time step $t-1$ and the input, followed by a point-wise non-linearity.

RNNs have problems with learning global dependencies in a sequence (Hochreiter, 1998; Bengio et al., 1994), and thus Long short-term memory (LSTM) architectures were invented (Hochreiter & Schmidhuber, 1997). An LSTM is a gated variant of the RNN where the cell at each time step contains an internal memory vector and three gates controlling what parts of the internal memory will be kept (the forget gate), what parts of the input that will be stored in the internal memory will be kept (the input gate), as well as what will be included in the output (the output gate). Another architecture was later proposed containing only two different gates, called the gated recurrent unit (GRU) (Cho et al., 2014).

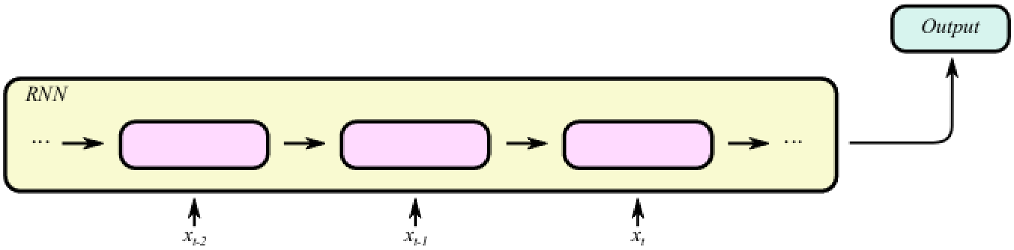

Figure 1: The proposed character based RNN model.

We train, evaluate and compare three different models: a feed-forward network, an RNN-LSTM network and an RNN-GRU network and show that it is possible to solve the problem of grammatical gender prediction using these models. The implemented networks consist of one hidden layer with 64 units, with a sigmoid activation function in the output layer. We train the models using binary cross-entropy loss. The dataset used to train and evaluate our models consists of 88,480 Swedish nouns acquired from SALDO (Borin, 2013), labelled with grammatical gender. In Swedish, about two thirds of nouns are utrum and one third neutrum. In our dataset, 71% of the words belong to the utrum class, and 29% to the neutrum. Further, the Swedish Academy notes that there are some common suffixes that strongly correlate with the grammatical gender. Words ending with ing, tion, het and ist are mostly associated with utrum and words ending with eri, skop and gram are associated with neutrum.

A batch size of 32 was used in our experiment for all models. Furthermore, all models were trained until no change was observed in the validation loss for 50 epochs. The models are evaluated using prediction accuracy, precision, recall and F1-score. The dataset was randomly split up into 60% training, 20% validation and 20% test. A second test set (13% of the dataset) was created from the first test set, where we removed all common suffixes that correlate highly with a gender. We can thus evaluate how well a model has learned to predict the grammatical gender of a word without suffixes to its disposal.

Figure 2: t-SNE visualisation of the output from the LSTM layer on test data.

In table 1 the results of the three models are presented on the test data set. We see that all models beat the 71% majority class baseline. The best performing model was the LSTM-RNN with a test classification accuracy of 95.15%. Further, this model achieved a 0.93 (0.89) and 0.96 (0.94) precision score, a 0.90 (0.85) and 0.97 (0.96) recall score and a 0.92 (0.87) and 0.97 (0.95) F1-score on the neutrum and utrum classes, respectively. The parenthesis values are calculated for the test set with removed words. In figure 2, a t-SNE visualisation of the representations of the LSTM is visualised and it is clear that the boundary in the middle between the classes is where most model mispredictions happen.

Model	Test set	Test set w/ removed words
Feed-forward	0.8548	0.8577
GRU	0.9492	0.9255
LSTM	**0.9515**	**0.9324**

Table 1: Test accuracy of the models.

Our results show that a simple feed-forward neural network achieves around 85% test accuracy, above the naive majority class baseline of 70%. A simple RNN-based model is able to predict grammatical gender in Swedish nouns (a lexical property that is considered difficult and does not adhere to rules other than in a few special cases) with a high test accuracy of 95%. Further, removing common suffixes from the test set does not impair the results markedly. Our results show that we can learn the necessary patterns only by looking at the sequence of characters in a given word, without using any context words.

References

Basirat, A. & Tang, M. (2018), Linguistic information in word embeddings. In *International Conference on Agens and Artificial Intellgience,* pages 492-513. Springer.

Bengio, Y. (1994). Learning long-term dependencies with gradient descent is difficult. *Neural Networks, IEEE Transactions on, 5(2):157-166.*

Borin, L. (2013). SALDO: a touch of yin to WordNet's yang. *Language Resources and Evaluation,* 47(4):1191-1211.

Cho, K., van Merrienboer, B., Gülçehre Ç., Bahdanau, D., Bougares, F., Schwenk, H & Bengio, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *EMNLP,* pages 1724-1734. ACL.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation,* 9(8):1735-1780.

Natase, V. & Popescu, M. (2009). "What's in a name?: in some languages, grammatical gender". *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing.* Volume 3-Volume 3. Association for Computational Linguistics, pp.1368-1377.

Rebbe, C., Gullberg, H & Ivan, N. (1954). *Svensk språklära.*

Teleman, U. & Hellberg, S. (1999). *Svenska akademiens grammatik.*