

## **Prediction of the next medication order to assist prescription verification by pharmacists in a health care center**

*Maxime Thibault, B.Pharm, M.Sc<sup>1</sup>, Maxime Bergeron, Pharm.D<sup>1</sup> and Denis Lebel, B.Pharm, M.Sc, FCHSP, FOPQ<sup>1</sup>*

*<sup>1</sup> Pharmacy Practice Research Unit (URPP), Department of Pharmacy, CHU Sainte-Justine, Montreal, Canada*

### **Background.**

In North American health care centers, medication orders are entered electronically or prescribed on paper. In most hospitals, pharmacists review all orders to ensure adequate medication usage and correct entry into the electronic medical record or pharmacy information system. As pharmacists move to a practice model involving direct patient care, the activity of drug order verification could benefit from a certain degree of automation to free up resources to perform clinical tasks. A potential application of machine learning would be to detect atypical orders and focus pharmacist attention on unusual patterns instead of verifying every single order. This could involve predicting drug orders and warning when actual orders deviate from those expected. Previous studies used conventional statistics or natural language processing alone and were either limited to selected drugs or limited in time. The objective of this project was to evaluate the feasibility of predicting the next prescribed drug using machine learning.

### **Methods.**

This study took place in a 500-bed mother-child, tertiary care university hospital center. Access to the institutional pharmacy database was authorized in accordance with local policies. A training set consisting of all orders from 2013 to 2017 inclusively was extracted, including anonymized patient encounter identifiers, drug identification, drug classes, and order start and end times. A test set from January to July 2018 was also extracted. Data was cleaned up and preprocessed to reconstruct the sequence of drug orders during every hospitalization so that each (label) order was associated with (features) the sequence of orders that preceded it (drug identifier, class and ordering department) and whether each preceding drug was still active or had been discontinued at the time of the order.

A neural network classifier was constructed using two inputs: 1) the sequence of word2vec embeddings of drug orders preceding the label and 2) a “bag of words” representation of active drugs and classes at the time of ordering encoded as a multi-hot vector. These features were fed into LSTM layers for the sequence and fully-connected layers for the multi-hot vector, followed by concatenation and fully-connected layers, with the next prescribed drug as the output. Word2vec embeddings were initially trained alone to maximize the accuracy on a list of analogies representing semantic relationships between drugs (e.g. a vector going from oral tablet to oral suspension for one drug). Then, multiple configurations and hyperparameters were tested to optimize performance of the neural network. All experiments were performed with 5-fold cross-validation with a time series split. Finally, the optimal model was tested against the 2018 set to validate performance on previously unseen data.

This model was deployed on the institutional intranet to show clinicians the rank in the predictions of the model associated with each actual order in a simplified fashion. Sampled patient profiles were used to construct a survey. Orders representing 20 clinical situations (routine order, simulated prescribing error, possible but rare combination, etc.) were presented with the profiles. Pharmacists were asked to rate how atypical they considered each order on a 4-point scale. The ranks of the model predictions were binned into 4 groups; cutoff ranks between groups were adjusted to maximize the accuracy of binned predictions as compared with pharmacist ratings.

### **Results.**

The 5-year training set contained 1,022,272 orders for 3145 drugs. The final word2vec embeddings showed a 77.1% (118/153) accuracy on the analogies. For prediction of the next order, baseline accuracy scores using a dummy classifier were, for the 2013-2017 set and 2018 set respectively, 4.5% and 5.1% for top 1, 23.6% and 21% for top 10, and 41.1% and 44.5% for top 30 accuracy. The final network configuration yielded a (mean  $\pm$  sd) cross-validation score on the 2013-2017 set of  $44.0 \pm 0.4$  % for top 1,  $68.2 \pm 1.3$ % for top 10 and  $78.8 \pm 1.3$ % for top 30 accuracy. The test set from 2018 contained 95,310 orders for 1843 drugs. 264 (0.3%) orders were discarded because of previously unseen labels. Results on this set were 44.4% for top 1 accuracy, 69.9% for top 10 accuracy, and 80.4% for top 30 accuracy, with a weighted average precision, recall and area under ROC curve of 0.415, 0.444 and 0.959 respectively. 18/35 pharmacists (51.4%) answered the survey. Inter-rater agreement of pharmacist ratings was poor with a Fleiss kappa of 0.283. After binning the model predictions, the accuracy of predictions compared with pharmacist ratings was 55.0% with a kappa of 0.338 and a weighted average precision and recall of 0.617 and 0.550.

### **Conclusion.**

Our prototype showed interesting performance for the prediction of the next prescribed drug. Comparison with pharmacist opinion proved difficult. Further research should focus on testing its usefulness in clinical practice.