| | |
|---|---|
| Name: | Kristin Feeney |
| Affiliation: | Deloitte Consulting LLP |
| Email: | krfeeney@deloitte.com |
| Presentation type (s): | **Poster** |

# Building Deep Learning Models with the OMOP CDM

**Kristin Feeney, MPH[1], Dennis Robert, MBBS/MD, MSST[1], Prerna Patil, MHIM[1], Sergey Charkin, MS[1], Dan Housman, BSc[1], Yuval Koren, MSc, [1] Jinlei Liu, MS[1]**
**[1]ConvergeHEALTH by Deloitte, Deloitte Consulting LLP, Newton, MA**

**Abstract**

*Advanced cognitive learning algorithms ("deep learning") are an emerging data science technique with many potential applications in the Life Sciences industry.[1,2] Using models such as convolutional neural networks (CNN) and recurrent neural networks (RNN) has the potential overcome limitations of conventional rule-based models that are unable of simulating the complexity of robust, interrelated data. To test this, we explored the use of CNNs and RNNs in large scale observational data transformed into the OMOP Common Data Model v5. Our experiment creates a framwork to evaluate the accuracy of a machine-driven view of disease etiology that triangulates the relationship between medications, procedures, diagnoses or labs preceding a diagnosis of interest. The "bag of features" in our model were created leveraging the OMOP concept mappings. We believe that the use of the OMOP CDM concept mappings allows for a robust feature generation process and provides control into adjustment for potential confounders. Our modeling exercise consists of comparing traditional modeling approaches (Random Forest, Logistic Regression) to deep learning models. Our output is an accelerator we call "Deep Miner" that combines proprietary data transformation and neural network algorithms to enable additional deep learning experimentation. Deep Miner is intended to be a methodology for enabling our Life Sciences clients with the methods to investigate their own deep learning hypothesis and contribute research to bolster the broader corpus. More research is needed to explore the utility of CNNs and RNNs for disease modeling.*

**References**

1. Xishuang Dong, Lijun Qian, Yi Guan, Lei Huang, Qiubin Yu, Jinfeng Yang, "A multiclass classification method based on deep learning for named entity recognition in electronic medical records", *Scientific Data Summit (NYSDS) 2016 New York*, pp. 1-10, 2016.
2. Jim O' Donoghue, Mark Roantree, Martin Van Boxtel, "A Configurable Deep Network for high-dimensional clinical trial data", *Neural Networks (IJCNN) 2015 International Joint Conference on*, pp. 1-8, 2015, ISSN 2161-4407.