The GatedTabTransformer: An Attention-Based Deep Learning Architecture for Tabular Modeling

Cholakov, Radostin (School: Model High School of Mathematics "Akademik Kiril Popov")

Some of the most common machine learning pipelines with real-world applications in the fields of healthcare, finance and marketing involve manipulation of tabular data. In recent years multiple new deep learning architectures have been able to set performance records mainly by applying concepts from other areas of machine learning such as natural language processing to tabular data. The current state-of-the-art solution for classification of such data is the TabTransformer model developed by Amazon in 2020. It incorporates a Transformer block with an attention mechanism to better track relationships between categorical features and makes use of a standard multilayer perceptron to output its final logits. I hypothesised that the model could be further enhanced with additional normalisation, linear projections and a gated MLP - a neural network introduced by Google. I propose modifications to the original TabTransformer outperforming it on binary classification tasks for three separate benchmark datasets with more than 1% area under receiver operating characteristic curve gains. To achieve this result the categorical data features are processed with attention, normalised, concatenated with continuous values and fed through multiple layers of channel projections and spatial gating units. I also evaluate the importance of specific hyperparameters during training. Multiple applications have been developed to showcase the potential of the architecture, including foetal health classification and grid management solutions for renewable energy plants. As a part of the project development we also implemented the GatedTabTransformer model in the Python library `tsai` for tabular modeling.

Awards Won:
Association for Computing Machinery: Second Award of $3,000