Paper for Discussion: Learning Feature Engineering for Classification
Fatemeh Nargesian, et al.
Presented by: Elizabeth Fitzgerald (efitzge5@nd.edu)
https://www.ijcai.org/Proceedings/2017/0352.pdf

The topic of this paper is that of feature engineering which is defined as, "... the task of improving predictive modeling performance on a dataset by transforming its feature space" [1]. In other words, feature engineering involves creating new features from existing values in a dataset by performing some transformation such as taking the ratio, arithmetic operations, logarithms, and other mathematical functions. However, creating new features is not always straightforward since knowing whether a constructed feature will improve a model's predictions/classification is not always clear.

The specific activity being studied is whether a method to automate the selection of useful features is possible without having to manually construct and test features. *Learning Feature Engineering* (LFE) is the proposed solution. LFE, when given a dataset, will output a recommended set of useful transformations; therefore automating the feature engineering process. LFE is a trained neural network using features (85k) extracted from 900 datasets.

The research question is not explicit, however the researchers aim to develop and evaluate a model based on the following, "the feature engineering problem is to find q best paradigms for constructing new features such that appending the new features to D maximizes the accuracy of L" [1]. Where D is the dataset, L is the particular model, and q is the number of useful features with the broader objective to do this automatically and efficiently.

The main problem being addressed is determining whether a method to perform automated feature engineering which reduces human involvement, is less computationally intensive, and is more accurate than current methods. The authors claim that it is possible to learn which transformations are useful, their model is capable of doing so without needing to apply them, and LFE is able to perform more efficiently than other methods. The evidence stems from their large use of features and datasets to train LFE on, which are then compared to other baseline methods. For statistical analysis, there is the measure of classification performance (F1 scores), compare the performance across datasets, evaluation of their neural network (LFE), and compare to baselines.

References:
[1] Nargesian, Fatemeh, et al. "Learning feature engineering for classification." *Ijcai*. Vol. 17. 2017.