# Classification of SAT Problem Instances by a Reusable MLP Neural Network

Márk Danisovszky, Zijian Győző Yang, Gábor Kusper

Eszterházy Károly University, Mathematics and Informatics Institute

{danisovszky.mark,yang.zijian.gyozo,kusper.gabor}@uni-eszterhazy.hu

## Abstract

Efficient SAT solving is critical in many practical applications.State-of-the-art SAT solvers can solve SAT problem with more than 100.000 variables and millions of clauses. They have usually more than 10 options, and one might combine them. We need those options to get better runtime. Our experience shows that we can get significant speed up if we use the suitable options. Our goal is to create a neural network based system, which can select the best option configuration for the input SAT instance. As a first step, we created a system that can predict the type of a SAT problem. Our system has two main tools: CNFStats and prediction tool. The CNFStats can compute several properties of the input SAT instance. For example, how many variables, clauses, unit, binary clause are in the input. How many horn clauses, so called black and white clauses are in the input. These properties generated by CNFStats are the input for the prediction tool. The prediction tool is a multi-layer perceptron neural network, which is able to do classification over SAT problem instances found on the SATLIB webpage, see: https://www.cs.ubc.ca/~hoos/SATLIB/benchm.html. The neural network is written in .NET from scratch. It is easy to configure by using a webpage. The user can give the number of layers, number of neurons, and the activation function, train it, test it and save it. In the prediction tool we use those 48 properties which is computed by the CNFStats tool. We used the half of the input problems from SATLIB to train the neural network. Classes were the same as in the webpage. The most successful neural network was using 2 hidden layers, 125 neurons in each hidden layer, the activation function was the sigmoid function (other activation functions did not perform well in these settings), 10.000 training runs with 0.033 learning rate. The accuracy was 95%. The second best neural network was the same but with 1 hidden layer and 250 neurons.