# Model card

## Model details

Chosen coding language: Python

Model date: 2021-03-26.

Model version: 1

Model type: ANN (Regression)

Author: Ivan Mitkov

## Intended use

Primary intended uses: The model predicts the next 7-days-ahead sales for all the restaurants locations of a certain chain in na rolling window, i.e. every day the model gives some forecasts for the next 7 working days.

Primary intended users: Logistic department and managers.

## Factors

The model power varies across the different locations.

## Metrics

Since the model deals with correctly predicting the sales of a chain, the so called "regression problem", the metrics, that are applied, originate from the domain of the regression models – MSE (Mean Squared Error) & MAE (Mean Absolute Error).

Results:

*Mean squared error:*

- Training set: 429982.97
- Validation set: 455845.24
- Test set: 2237858.27

*Mean absolute error:*

- Training set: 418.91
- Validation set: 426.35
- Test set: 972.87

*Mean squared error per days ahead (test data):*

- 1: 2.002642e+06
- 2: 2.077126e+06

- 3: 2.255014e+06
- 4: 2.409453e+06
- 5: 2.350739e+06
- 6: 2.268718e+06
- 7: 2.455643e+06

## Data

The training (& validation) data set consists of samples for all chain locations from 2017-12-01 until 2010-11-30.

As a training data set are used 80% randomly selected rows from this period of time, while the other 20% serve as a validation data set. It should be stressed, that the training data is randomly chosen in order to keep the time variance incorporated in it. The validation data has the main goal to stop the model from "learning" the finest details and pattern, that could lead to overfitting.

In order to capture the time fluctuation of the sales, some new features are added to the data set. By this kind of feature engineering, for example, the model considers the most recent sales for a specific day or the previous day sales ect.

Before feeding the data to the ML algorithm, the variables have been also normalized, this aims a better performing of the model.

## Quantitative Analysis

Since the chain is predicting few sequences (predictions for different locations) and the sequence are of different length, the most suitable algorithm, that could be deliver results for the first iteration, is a smple form of Artifictial Neural Network with 2 hidden layers.

The graph bellow gives an overview how the model "learned" during the training phase.

As a focus of future work, one should mark the fine-tuning of the model, finding an appropriative way for developing of RNNs or even ensemble of algorithms and also deploying the model in a more user friendly way..