Разбор статьи "Predicting the direction of stock markets using optimized neural networks with Google Trends" Hongping Hu, Shuhua Zhang, Haiyan Wang

# Постановка задачи

Предсказание направления рынков ценных бумаг с помощью оптимизированных нейронных сетей и используя Google Trends.

# Данные

Используются дневные данные (цены открытия, верха, низа и закрытия, объём торгов) двух индексов: S&P500 (Standard & Poor 500) и DJIA (Dow Jones Industrial Average) с 1 января 2010 по 16 июня 2017, полученные с сайта finance.yahoo.com. Также используются ежедневные метрики (степени релевантности поисковых запросов) по ключевым словам "S&P 500" и "DJIA" в Google Trends.

# Предложенный метод решения

Улучшается алгоритм "Sine Cosine Algorithm" (SCA), введением дополнительного коэффициента. Далее полученный алгоритм "Improved Sine Cosine Algorithm" (ISCA) используется для определение оптимальных параметров нейронной сети с обратным распространением ошибки "Back propagation neural network" (BPNN) и создания модели ISCA-BPNN. В ISCA p агентов определяют глобальный оптимум на T итерациях. Каждый агент представляет собой вектор, который необходимо оптимизировать. Этот вектор содержит параметры BPNN: веса w и нейроны смещения b. Сопоставляя каждого агента с BPNN, получается предсказанное значение. Далее минимизируется функция приспособленности (fitness function), которая представляет собой метод наименьших квадратов (МНК) между предсказанными значениями и наблюдаемыми (цены открытия), усреднённая на тестовую выборку.

# Актуальность / обоснование метода

С целью обойти ограничения градиентного поиска (gradient search) в методе обратного распространения ошибки (BP) используется нейронная сеть с обратным распространением ошибки (back propagation neural networks) с алгоритмом ISCA для улучшения предсказания направления движения рынков.

# С чем сравнивают

Все измерения проводят по 10 раз, выбирая лучший результат модели и средний. Модель ISCA-BPNN сравнивается с самой BPNN и вариациями BPNN, использующими другие алгоритмы для определения весов (GWO, PSO, WOA, SCA). У моделей с алгоритмами PSO и ISCA варьируют параметры и . Затем рассматриваются методы, исследованные в других работах: Генетический алгоритм с признаковой дискретизацией (GA feature discretization), классификатор (Classification model), метод опорных веторов (SVM), нейронная сеть с обратным распространением ошибки (BPNN), а также гибрид генетического алгоритма и нейронной сети (GA-ANN hybrid model). Также модели сравнивается с теми же, но без использования информации Google Trends. Результаты рассматриваемых моделей из других статей не могут быть показательными, потому что измерялись на других рынках и в другой промежуток времени.

# Результаты

Предложенная модель ISCA-BPNN с различными параметрами стабильно давала лучшие результаты (хоть и незначительно) на обоих индексах S&P500 и DJIA как с использованием данных Google Trends так и без.

# Доступность данных

Данные находятся в открытом доступе. Для цен индексов используется finance.yahoo.com. Для поисковой метрики - trends.google.com (возможно выставление геолокации на США для более релевантных данных)

# Доступность реализации

В статье не опубликованы источники на открытый код проекта. Сложность самого проект умеренно-сложная.

Плюсы / минусы статьи по вашему мнению  
Результаты представлены вероятностью правильного предсказание дневного направления. На практике вероятность не показательна, так как даже модель, правильно предсказывающая направление с вероятностью 0.9, может быть убыточна, если убытки сильно превышают выигрыши.

# Идеи для дальнейших исследований (указанные в статье / предложенные вами)

Исследователями планируется использование оптимизированных нейронных сетей для предсказания цен и торговых объёмов вместо направлений. Также планируется использовать больше ключевых слов в Google Trends, а также установление влияния Google Trends на цены акций и объёмы сделок.

Ахияров Валентин, МФТИ

P.S. Оформление разбора в TeX'e планируется в субботу вечером (13 Oct 2018)