

## Задача 1

Название: Прогнозирование направления движения цены биржевых инструментов по новостному потоку.

Задача: Построить и исследовать модель прогнозирования направления движения цены. Задано множество новостей  $S$  и множество временных меток  $T$ , соответствующих времени публикации новостей из  $S$ . 2. Временной ряд  $P$ , соответствующий значению цены биржевого инструмента, и временной ряд  $V$ , соответствующий объему продаж по данному инструменту, за период времени  $T'$ . 3. Множество  $T$  является подмножеством периода времени  $T'$ . 4. Временные отрезки  $w=[w_0, w_1]$ ,  $l=[l_0, l_1]$ ,  $d=[d_0, d_1]$ , где  $w_0 < w_1=l_0 < l_1=d_0 < d_1$ . Требуется спрогнозировать направление движения цены биржевого инструмента в момент времени  $t=d_0$  по новостям, вышедшим в период  $w$ .

Данные:

Финансовые данные: данные о котировках (с интервалом в один тик) нескольких финансовых инструментов (GAZP, SBER, VTBR, LKOH) за 2 квартал 2017 года с сайта Finam.ru; для каждой точки ряда известны дата, время, цена и объем.

Текстовые данные: экономические новости за 2 квартал 2017 года от компании Форексис; каждая новость является отдельным html файлом.

Литература:

Usmanova K.R., Kudiyarov S.P., Martyshkin R.V., Zamkovoy A.A., Strijov V.V. Analysis of relationships between indicators in forecasting cargo transportation // Systems and Means of Informatics, 2018, 28(3).

Kuznetsov M.P., Motrenko A.P., Kuznetsova M.V., Strijov V.V. Methods for intrinsic plagiarism detection and author diarization // Working Notes of CLEF, 2016, 1609 : 912-919.

Айсина Роза Мунеровна, Тематическое моделирование финансовых потоков корпоративных клиентов банка по транзакционным данным, выпускная квалификационная работа.

Lee, Heeyoung, et al. "On the Importance of Text Analysis for Stock Price Prediction." LREC. 2014.

Базовый алгоритм: Метод, использованный в статье (4).

Решение: Использование тематического моделирования (ARTM) и локальных аппроксимирующих моделей для перевода последовательности текстов, соответствующих различным временным меткам, в единое признаковое описание. Критерий качества: F1-score, ROC AUC, прибыльность используемой стратегии.

Новизна: Для обоснования связи временных рядов предлагается метод Сходящегося перекрестного отображения.

Авторы: В.В. Стрижов (эксперт), К.В. Воронцов (эксперт), Иван Запутляев (консультант)