Аннотация:

В работе рассматривается задача классификации направления движения временных рядов.

Классификация производится с помощью анализа признаков из отчётов 8-K, которые компании обязаны заполнять

при значительных событиях, таких как банкротство, выбор совета директоров и пр.

Рассматривается несколько моделей классификации.

В одних используются только признаки из отчётов, 1-граммы которых встречающиеся более 10 раз.

В других к предыдущему этапу применяется неотрицательная матричная факторизация (NMF).

И в последней, ансамбле, объединяются предыдущие подходы путём голосования большинства.

В качестве прикладной задачи рассматривается задача распознавания направления движения акций по новостям, выраженных

отчётами 8-K.

Модели классификации, исследованные в этой

работе, сравнивается в точности и статистической значимости с простыми моделями, использующими только прогнозируемый показатель доход на акцию

или использующую только финансовые показатели.

Ключевые слова: метрическая классификация, анализ текстов, классификация

временных рядов, новостной поток