Отзыв

на магистерскую диссертацию Гао Тянь

«Разработка новых методов анализа настроений, анализа текста в социальных сетях для прогнозирования предпочтения пользователей»

Решение задачи, которую магистрант поставил перед собой, требует знаний в области программирования, математики, лингвистики, психологии. Необходимо также хорошее знакомство с различными программными средствами и умение их использовать.

Представленная работа демонстрирует, что магистрант отлично владеет всеми необходимыми знаниями и умениями.

Работа состоит из обширного введения, шести глав, выводов и заключения.

Автор рассматривает два принципа определения эмоциональной окраски текстов. Первый – по ключевым словам, второй – по результатам машинного обучения.

Обсуждаются способы получения данных как для анализа настроений, так и для обучения нейронных сетей. Стратегии работы веб-краулеров

Большое внимание автор уделяет методам разделения слов для английского, русского и китайского языков. Описываются также методы очистки данных (борьба с шумами, дупликацией и другими проблемами).

Подробно описываются методы векторизации слов: модель Word2Vec, ее развитие GloVe и программный пакет от Google – Bret.

Значительное внимание уделено описанию конволюционной нейронной сети – CNN.

Немного необычен стиль изложения. Описание современных методов и собственных разработок ведется параллельно. Мне это не кажется недостатком, так как позволяет, на мой взгляд, лучше увидеть все этапы работы автора и их связь с современными мировыми достижениями.

Изложение ведется на хорошем русском языке.

Автор, Гао Тянь, является сформировавшимся высококвалифицированным специалистом, его диссертация, после небольшой доработки, может и должна быть опубликована.

Я оцениваю работу на отлично.

Научный руководитель канд. физ.-мат. наук, доцент

 /Сергеев С.Л./