

Отзыв

Научного руководителя

на выпускную квалификационную работу Гарипова Эмиля Ильдаровича на тему:
«Автоматические методы анализа социологических данных», представленную для
защиты в 2016 году.

Работа Эмиля Ильдаровича относится к важному и актуальному направлению современных исследований, а именно – автоматическому анализу данных, используемых при социологических исследованиях. За последние несколько лет можно наблюдать большой рост в использовании социальных сетей и платформ для микроблоггинга; размещаемая информация, в первую очередь текстовая, составляет большую базу для социологических исследований. Обработка текстов вручную с целью извлечения полезной информации требует слишком много времени и человеческих ресурсов. Одной из задач социологических исследования является анализ тональности текстов, публикуемых в социальных сетях, и ее высококачественное автоматическое решение является очень актуальным. Существующие решения автоматического определения тональности текстов требуют анализа, проверки на различных наборах данных и адаптации.

Конкретно, в работе решается задача анализа тональности коротких текстов, размещенных в социальной сети Twitter. В рамках работы Эмиль Ильдарович подготовил обзор методов автоматического определения тональности текстов; провел сбор и предобработку данных для обучения и тестирования моделей; провел экспериментальное сравнение методов обучения с учителем для решения задачи классификации по тональности текстов на русском языке: наивный байесовский классификатор и метод опорных векторов при использовании различных способов извлечения признаков.

Особенностью работы является сопоставление моделей на нескольких наборах данных имеющих разную природу, что позволяет не только получить абсолютные оценки качества моделей, но и проанализировать особенности их поведения, например, оценить возможность применения модели, натренированной на коллекции данных из одной предметной области, в новом окружении.

Вследствие плохого планирования работы проведение экспериментов, так и подготовка текста работы проходили в условиях крайней спешки, что не могло не сказаться на полноте проведенного экспериментального исследования и качестве представленного текста.

Работа Э.И. Гарипова заслуживает оценку «Хорошо».

ст. преп. каф. ИАС СПбГУ



А.С. Ярыгина