МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Емельянова Анастасия Вадимовна

«Процедура семантического отображения номинативного компонента текста в терминах WordNet-словарей»

Профиль: «Компьютерная лингвистика и интеллектуальные технологии»

Программа: «Компьютерная лингвистика и интеллектуальные технологии»

Научный руководитель: к. ф. н., доц. Азарова И. В.

Задача описания смысла текста на естественном языке является одной из краеугольных проблем теоретической и прикладной лингвистики. В данной работе, согласно теории Ш. Балли, под интерпретацией текста понимается выявление диктума и модуса высказывания. Материалом послужили тексты отзывов на цифровые фотоаппараты пользователей сервиса «Яндекс. Маркет». Ограниченность предметной области позволяет более четко сформулировать процесс интерпретации. Поскольку тексты отзывов являются статическими (в терминологии А. С. Штерн), то есть описательными, то в качестве основного объекта анализа высказывания выступает номинативный компонент, а не структура пропозиции.

В исследуемых текстах диктум выражен понятиями предметной области «Цифровые фотоаппараты», поскольку именно фотокамеры, их части, функции и характеристики являются объектами описания. Модус – оценка говорящего, мнение о продукте в целом и его частях по отдельности.

В результате исследования построена логико-понятийная схема предметной области и разработан алгоритм извлечения компонентов объективной составляющей значения текста, а также представлен алгоритм оценки общей тональности текста по оценочным лексическим маркерам, построенный с применением нейронной сети. Кроме того, описан механизм оценки отдельных компонентов логико-понятийной схемы, в основе которого лежит анализ комбинаторных характеристик тематических и тональных маркеров, присутствующих в тексте.

Результаты данной работы могут применяться при автоматическом поиске и статистической обработке текстов отзывов при проведении маркетинговых исследований.