

**ПРИМЕНЕНИЕ НЕЙРОСЕТЕВОЙ РЕАЛИЗАЦИИ
EASY-FIRST ПОДХОДА В ЗАДАЧЕ РАСПОЗНАВАНИЯ
ИМЕНОВАННЫХ СУЩНОСТЕЙ ДЛЯ РУССКОГО
ЯЗЫКА**

Петров Максим Андреевич

Аспирант

*Лаборатория Нейронных систем и глубокого обучения, МФТИ, Москва,
Россия*

E-mail: maksimallist@gmail.com

Алгоритмы машинного обучения, основанные на архитектурах искусственных нейронных сетей, в настоящее время демонстрируют передовые результаты в задачах обработки естественного языка. Распознавание именованных сущностей вызывает повышенный интерес [2-3], так как существующие на данный момент решения обладают недостаточно высоким качеством работы, и одновременно алгоритмы способные решать данную задачу на высоком уровне, чрезвычайно востребованны в IT отрасли, с точки зрения реальных приложений.

В работе [1] была впервые предложена полностью дифференцируемая нейросетевая архитектура, позволяющая реализовать easy-first подход при решении различных задач обработки естественного языка. Что позволило обучать подобные модели с помощью градиентных методов обучения.

В данной работе предложенный подход применялся для решения задачи распознавания именованных сущностей на русском языке. Эксперименты проводились на корпусе FactRuEval 2016. Было показано, что применение данного подхода, совместно с современными нейросетевыми архитектурами, улучшает качество работы алгоритмов по всем метрикам, принятым для измерения качества систем распознавания именованных сущностей (accuracy, F-1).

Литература

1. Martins A. F. T., Kreutzer J. Learning What's Easy: Fully Differentiable Neural Easy-First Taggers // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. – 2017. – С. 349-362.
2. Sysoev A. A., Andrianov I. A.: Named Entity Recognition in Russian: the Power of Wiki-Based Approach. dialog-21.ru
3. Reproducing Russian NER Baseline Quality without Additional Data. In proceedings of the 3rd International Workshop on

ConceptDiscovery in Unstructured Data, Moscow, Russia, 54 – 59
(2016)