

Т. С. Петунина, А. В. Бурмакова, Н. А. Жилияк
(БГТУ, Минск)

ГЕНЕРАЦИЯ ОСМЫСЛЕННЫХ УНИКАЛЬНЫХ ТЕКСТОВ

Одна из главных проблем с которой сталкиваются SEO-специалисты – это написание тематических, уникальных и осмысленных текстов, необходимых для продвижения сайтов в поисковых системах. Заказ написания уникальных текстов является довольно затратным делом, что привело к идее об автоматической генерации текста.

В этой статье речь пойдёт не о генерации простого бессмысленного набора слов, а о создании осмысленных и полезных текстов. Самой главной проблемой является автоматизация генерации текста. Не редко высказывается мнение, что автоматически эту задачу решить нельзя, но на практике это убеждение уже устарело.

В качестве задачи, будет рассматриваться вопрос генерации уникального описания товара на торговой площадке. Имея пару описаний товара, полученных из официальных источников, необходимо создать краткое и уникальное описание товара соответствующее действительности.

Например, необходимо написать уникальное описание конкретного смартфона, какую информацию следует включить в текст? Необходимо определить производителя, модель, оперативную память смартфона, разрешение экрана и т. д., так же дополнительные его характеристики, но что из этого действительно является важной информацией для пользователя? Самый простой способ – это анализ

по словарю вхождений слов, таких как «память», «камера», «ёмкость батареи», «операционная система» и т. п. Более точный способ выделения аспектов и их оценок может опираться на обучаемую систему выделения информации из текста.

Решить данную проблему можно следующими способами:

– использование шаблонов. Используются заранее заготовленные предложения, производится анализ признаков и составляются соответствующие предложения;

– использование формальных грамматик. Используется набор правил с помощью которого текст генерируется случайным образом;

– использование нейронных сетей. Используются системы, которые учатся генерировать новые тексты на существующих текстах.