

РАЗРАБОТКА АЛГОРИТМОВ ВАЛИДАЦИИ И УТВЕРЖДЕНИЯ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ, УЧЕБНЫХ ПЛАНОВ

А. Арзуманян

(Университет ИТМО, Санкт-Петербург)

Научный руководитель – к.т.н. М. В. Хлопотов

(Университет ИТМО, Санкт-Петербург)

Введение. Тенденции современного общества требуют новых подходов к организации образовательного процесса. Индивидуализация, как один из новых подходов, позволит студентам более эффективно овладеть наиболее необходимыми компетенциями в профессиональной сфере, открывая возможности выбора, учитывая их личную заинтересованность. Перевод учебных дисциплин в векторное пространство позволит применять алгоритмы машинного обучения, находить зависимости между учебными сущностями, а также создавать оптимальные рекомендации выбора дисциплин и построения индивидуальной траектории.

Целью работы является исследование методов перевода учебных дисциплин в векторное пространство.

Задачи:

- Проанализировать существующие методы векторизации;
- Определить наиболее подходящий метод перевода дисциплин в векторное пространство, который будет учитывать несколько уровней связей дисциплин;

Векторное представление — общее название для различных подходов к моделированию языка и обучению представлений в обработке естественного языка, направленных на сопоставление словам из некоторого словаря векторов небольшой размерности.

В ходе работы были проанализированы существующие методы векторизации текстов. Основными сущностями при построении векторных представлений текстов на естественном языке являются токен, документ, корпус и словарь. Помимо анализа токенов, часто рассматриваются n-граммы – последовательности из n подряд идущих токенов. Однако данные подходы имеют ряд недостатков, одним из которых является постоянно изменяющаяся размерность полученных векторов. Наиболее популярный способ, который решает данные проблемы – это использование эмбедингов. Для того, чтобы обучить нейросеть, нужно определить какие-то признаки — то есть, ряд неких параметров — отдельными наборами. Каждому набору должна соответствовать метка. В нашем случае меткой — будет дисциплина, а набором параметров — ее контекст. Далее требуется определить, что будет являться контекстом учебной дисциплины.

Выводы. На основании проведенного анализа существующих методов векторизации был определен наиболее подходящий метод перевода учебных дисциплин в векторное пространство, который бы учитывал как междисциплинарные связи, так и связей с учебными планами, направлениями, онлайн-курсами и т.д.

Автор:

Арзуманян А.

Научный руководитель:

Хлопотов М.В.