

Разработка инструментов интерактивной визуализации для веб-сервиса работы с байесовскими сетями web-BAMT

Актуальность

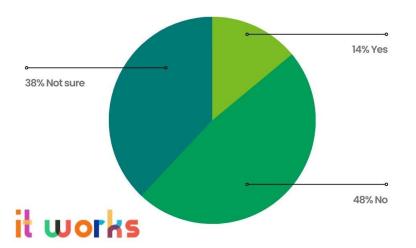


Люди не имеют адекватного представления о работе ИИ и не доверяют результатам работы



Необходимо разработать инструмент, с помощью которого можно изобразить процесс принятия решений со стороны ИИ

Для этого были выбраны Байесовские сети, как лучшая интерпретация ИИ из-за ее простого графового представления Would you trust a brand that uses an artificially intelligent algorithm and not a person to write content for their website?



Цели и задачи



Рассмотреть причины слабого доверия к системам ИИ



Проанализировать способы (алгоритмы) пригодные для интерпретации комплексных систем и представления их в удобной для пользователя форме

Выделить роль Байесовских сетей в концепции объяснимого ИИ (XAI)

Провести обзор существующих программных решений для визуализации БС

LITMO



XAI - концепция разработки моделей ИИ с фокусом на том, чтобы систему возможно было представить конечному пользователю, не обладающему экспертных знаний в машинном обучении

Байесовская сеть - вероятностная модель в виде ориентированного ациклического графа, представляющая собой множество переменных и их вероятностных зависимостей

Интерпретация модели - представление процесса работы, например структуры модели, демонстрирующее переход от получения входных данных до результата

Литературный обзор



Основная причина отсутствия доверия - непрозрачный механизм работы ИИ



Если и отображать схему внутренностей модели, ее нужно уменьшить для удобного восприятия. Разобраны методы РСА, Sammon Mapping, LIME

Из существующих программ для визуализации БС сделан обзор на Bayes server, BNView и ShinyBN

Литературный обзор



Название	Отображение графовой структуры	Ретроспектива процесса обучения	Сводка о результатах работы	Веб-приложение
Bayes server	Нет	Нет	Можно задавать вероятностные запросы и генерировать блоки со статистикой	Есть
BNViewer	Есть	Есть	Дополнительно с помощью других функций R	Нет
ShinyBN	Есть	Нет	Встроенное в АРІ вероятностных запросов на основе библиотеки gRain	Есть





Вывод



Ознакомился с новой концепцией XAI и теоретической базой для Байесовских сетей



Проведен обзор программ для визуализации БС

План работы



Ознакомиться со структурой проекта <u>web-BAMT</u>



Учитывая опыт рассмотренных фреймворков создать визуализацию, удовлетворяющую всем условиям из таблицы

VITMO

