

image not found or type unknown



## Вступление

Надеюсь, это эссе предполагает свободный стиль. Я пока точно еще не сформулировал сам для себя тему НКР, подобного рода работы и должны помочь в этом выборе. Ранее я изучал точные науки и прикладную информатику в экономике, сейчас точно уверен, что мне хочется связать свою работу с машинным обучением. Поэтому название сейчас достаточно общее, в процессе размышлений и написания подобных работ оно, предполагается мне, будет становиться все более и более узконаправленным.

## Основная часть

Машинное обучение (Machine Learning) -- это раздел прикладной математики, объединяющий методы математической статистики, оптимизации и искусственного интеллекта. Машинное обучение исследует методы построения алгоритмов, способных обучаться на основании эмпирических данных.

Частью машинного обучения является интеллектуальный анализ данных - совокупность подходов, методов и алгоритмов, которые используются для автоматического получения каких-либо нетривиальных выводов на основе эмпирических данных, в том числе сигналов, изображений, текстов, и т. д.

Машинное обучение необходимо для автоматизации решения сложных профессиональных задач. В силу универсальности используемых алгоритмов машинное обучение применимо в самых разных областях деятельности. Например, оно используется для распознавания речи, жестов, рукописного текста, в сфере медицинской диагностики и биржевого анализа, для борьбы со спамом, осуществления финансового контроля и во многих других сферах. Сфера применений машинного обучения постоянно расширяется, так как стремительная информатизация общества приводит к накоплению огромных объёмов эмпирических данных в науке, производстве, социальных областях. А возникающие при этом задачи прогнозирования, управления и принятия решений часто требуют быстрого анализа этих объёмов.

В этом случае методы машинного обучения оказываются не только применимыми, но и зачастую более эффективными в условиях, когда человеческий мозг не в силах

проанализировать огромный объем данных достаточно быстро. Огромные вычислительные мощности современных компьютеров могут оказать человеку в этом существенную помощь.

Сейчас наиболее интересным мне представляется интеллектуальный анализ данных (Data Mining) экономической информации. Семантический анализ текстов социальных сетей. Например, нас могут интересовать конкретные продукты или бренды.

Актуальность:

На сегодняшний день, в век информационных технологий, общество стремится быть социальным. Сейчас человек взаимодействует, так или иначе, с сотнями, а то и тысячами людей в день иногда даже не выходя из дома. Миллионы сообщений, постов в социальных сетях и прочих ресурсах. Все это хранит знания о настроении общества, рейтинге, популярности товара или бренда, общественно значимого события, остается только получить эти знания. А это на сегодняшний день не возможно без машинного обучения.

Первое преимущество – автоматизация. С помощью машинного обучения можно автоматизировать процессы, на первый взгляд плохо поддающиеся автоматизации.

Второе преимущество - алгоритмы машинного обучения, позволят выявить в данных скрытые взаимосвязи, которые простым аналитическим путем будут не видны.

Так же к актуальности можно отнести и причину популярности этой области. Методы нейронных сетей и машинного обучения были известны давно, впервые заговорили об этом еще в 1950 годах. Но только с недавнего времени вычислительные мощности позволили на практике применять такие подходы.

В целом в ходе написания работы хочется рассмотреть несколько задач. Оценить объемы потенциально полезной информации, находящейся в открытом доступе. Оценить затраты на сбор и обработку, поддержание актуальности тех или иных выводов. Ведь с течением времени «сводка» должна постоянно актуализироваться. Рассмотреть разные варианты использования полученных знаний. Возможный экономический эффект от них.

## **Заключение**

Машинное обучение - одна из ключевых информационных технологий будущего; наиболее успешное направление искусственного интеллекта, вытеснившее экспертные системы и инженерию знаний; математическое моделирование, когда данных много, знаний мало; тысячи алгоритмов; около 100 000 научных публикаций в год.

Если нет данных, то нет и машинного обучения. Сейчас в социальных сетях ежедневно публикуется миллионы пользовательских записей в открытом доступе. Это не только большой простор данных, но и отражение актуального настроения общества по той или иной теме. Остается только получить из данных - знания.