Идеи за курсови проекти по ИИ

## Идеи за проекти

1. **Text summarization (long documents; scientific papers; legal documents; …).**
2. **Federated Learning**
   1. 1. Да се имплементира оптимизатор на Адам, без да се използва готова библиотека на високо ниво като Tensorflow и PyTorch (from scratch) и да се покаже сходимост на алгоритъма.
   2. 2. Да се направи класификатор на изображения, използвайки данните от множеството CIFAR-10 и модел на невронна мрежа, създадени с библиотеките PyTorch или Tensorflow.
   3. 3. Да се направи класификатор на изображения на пътни знаци, който да работи с множеството GTSRB (German Traffic Sign Recognition Dataset) и да се провери неговата точност.
3. Explainable AI (ML)
4. **Оцветяване на снимки на София.**
5. Система подпомагаща избягването на чуждиците в българския език. Да се разработи система, която да подпомага в избора на подходяща дума на български, например: написано е „база“ системата предлага да я смени с „основа“, визия-виждане и т.н.
6. Разговорни агенти chatboot.
   1. Обучават се от данни за подобни диалози
   2. Обучават се от „учебник“ и „книги“
   3. С правила и/или в комбинация с по-горните.
   4. За конкретно приложение.
7. Разбиране на текст – агента „чете“ и „разбира“ текста.
8. Автоматизирана проверка и оценка на тестове в изпитни тестове.
9. Автоматизирано генериране на изпитни тестове от учебник.
10. Анализ на не комерсиални данни/текстове: публични данни, от е-правителство, научни данни, образование, здравеопазване и др.

[**https://opendata.government.bg/en/**](https://opendata.government.bg/en/)

1. Правене на разюме на книга (summarization). Тук задача се усложнява от търсен представяне на сюжета на книгата, преставане жанра, тропи (<https://bg.wikipedia.org/wiki/%D0%A2%D1%80%D0%BE%D0%BF>) …
2. Анализ на емоции в текст

<http://changingminds.org/explanations/emotions/basic%20emotions.htm>

От линк по-горе примерни задачи за анализ/класификация/и други на текстове:

* Изследване дали дървото на основните емоционални е се потвърждава от статистика на съ-срещанията в корпус от документи от дадена категория – например: книги, статии web и др.
* Помагат ли емоциите като характеристики за класифициране на книги, новини и др.
* Множество от относително ортогонални емоционални оси от обединението от основните емоции. Колко и кои?
* Проектира не останалите прилагателните имена в пространството на емоционалните оси.

<http://changingminds.org/explanations/emotions/five_dimensions_emotions.htm>

|  |  |
| --- | --- |
| **Positive emotions** | **Negative emotions** |
| Happiness, liking, respect, hope, contentment | Sadness, anger, fear, shame, disgust |

Превод на български с google translate

|  |  |
| --- | --- |
| **Положителни емоции** | **Отрицателните емоции** |
| Щастие, обич, уважение, надежда, доволство | Тъга, гняв, страх, срам, отвращение |

Примерни задачи от линк по-горе:

* Може ли емоциите да помогнат за определяне на полярност на текст.
* Другите емоции (от таблиците в предишния линк) могат ли да се класифицират като положителни и отрицателни.
* До колко горните оси са противоположни: (щастие – тъга (нещастие)?) (Обич – гняв (омраза)?) … И независими?
* **Важна задача: Горните и други изследвания за Български език.**

Предишни изследвания:

* Дипломанта работа на Тодор Новаков
* за български:   
  <http://sci-gems.math.bas.bg/jspui/bitstream/10525/2386/1/ERIS2014-book-p13.pdf>

1. **Обработка на медицински текстове за по-ефективно търсен (в сътрудничество с** доц. Бойчева) – виж приложен подробен списък
2. Приложение на DM методи за откриване на зависимости в данни с използване на допълнителни знания (background knowledge, ontologies, ...)
3. Visualization
   1. 1. Система за визуализация на многомерни данни - вектори, клъстери, и т.н.
4. Разнообразяване на препоръчаната информация – например показване и на алтернативно мнение за дадено събитие теория и т.н.
5. Обществени поръчки

Има публично достъпна информация за обществени поръчки.

Данни:

<http://opendata.government.bg/organization/www-aop-bg>

История на обществените поръчки от 2011-2015 ~ 100к

Задача:

1. Да “предсказваме” участниците
2. Победителят в дадена обществена поръчка

**Идеи за проекти (“Изкуствен интелект”)**

1. Инструмент за визуализиране на алгоритми. Идеята е да се направи инструмент, който да предостави удобен интерфейс, със следната функционалност:
   1. Лесно въвеждане на конкретен граф или част от характеристиките му (да поддържа свързани и несвързани графи, насочени и ненасочени, с тегла, с евристики и т.н.).
   2. Да се визуализира пълно обхождане и намиране на целево състояние.
   3. Да включва всички алгоритми, които сме учили и да предоставя лесно добавяне на нови.
   4. Да дава възможност за визуализиране на структурата от данни за обхождане на графа (какви обекти има в нея). Също така да поддържа и визуализация на структура за посетени върхове.
   5. Да се визуализира къде се намираме в примерен псевдокод по време на всяко действие.
   6. Да е пригоден за лесно качване на сайт/сървър.
2. Инструмент за автоматично създаване на програма за ФМИ използвайки CSP.
3. Оцветяване на изображения (стари снимки на София).
4. Edge Detection с невронни мрежи
5. Класификация на ICD-10 кодове въз основа йерархия.
6. Система за разпознаване на чуждици.
7. Система за коригиране на медицински термини.
8. Система за откриване и трансформиране на дати и съкращения в медицински текстове (“ст II” да стане “степен 2”)
9. Създаване на чат бот, който да отговаря на административни въпроси свързани с ФМИ.
10. Създаване на стратегия за федерирано обучение (Federated Learning).
11. Търсачка за сайта на ФМИ.
12. Клъстеризиране на файлове и документи.
    1. На изображения (снимки).
    2. На текстове (Word, PDF, txt).
    3. На аудио и видео.