

АЛЕКСАНДРА ЗИМИНА: ТОЧНОСТЬ ДИФФЕРЕНЦИАЛЬНОЙ ДИАГНОСТИКИ – 97%

14 октября 2024

Студентка 2 курса ФБИУКСа направления «Бизнес-информатика» Александра Зимина уже серьезно занимается научной работой, а именно – обучает нейронные сети для разных исследований, от медицинских до социологических. Пресс-служба МИФИ побеседовала с молодым ученым о ее проектах и интересах.



Александра Зимина, студентка 2 курса ФБИУКС, Б 23 902

- Как ты выбрала эту тему – нейронные сети и медицина?

- Когда я училась в 10 классе нашего лицея № 1511, IT-классы и проектную практику у нас курировал Алексей Егоров, под его руководством я выбрала тему – решила заняться онкологией и взяла проект «Диагностика онкологии с помощью нейронных сетей». А потом сфокусировалась уже на диагностике онкологии легких по КТ с помощью нейронных сетей. То есть я занимаюсь этой темой уже несколько лет, и моя точность в диагностике достигла 97%. С этой работой я выступала в апреле 2024 года на Международной научной школе-конференции НИЯУ МИФИ по профилю «Цифровые технологии и их применение» и стала победительницей. Сейчас готовлю публикацию для «Вестника МИФИ».

- В чем новизна твоего подхода?

- Это дифференциальная диагностика. С онкологией легких при постановке диагноза врачи сейчас в основном опираются на бинарную классификацию – либо есть заболевание, либо нет. А узнать какая именно онкология – сложно, но ведь от вида онкологии зависит лечение.

- Каков алгоритм работы врача, пользующегося твоим приложением?

- Врач загружает снимок КТ в специальную программу, нейронная сеть обрабатывает его и выдает результат – какой это вид опухоли: это может быть аденокарцинома, крупноклеточная карцинома, плоскоклеточный рак легкого, последнее – это широко распространенное онкологическое заболевание, на его долю приходится до 40-50 % всех случаев возникновения злокачественных новообразований, возникающих в бронхолегочной системе. Или – это не опухоль. По дифференциальной диагностике сейчас пока мало работ, это вообще проблемная область. Понятно, что даже при очень высоком результате за 90% нельзя полагаться только на ИИ, но такой инструмент в помощь врачам очень нужен. Онкология легких – второе по распространенности раковое заболевание, после онкологии молочной железы. На третьем месте – онкология толстой и прямой кишки.



- Ты делала в школе проекты по этой медицинской теме, а решила поступать в МИФИ на бизнес-информатику?

- Я люблю заниматься чем-то сразу в нескольких областях, на стыке разных дисциплин. Бизнес-информатика – это стык экономики и бизнеса, а мой проект – на стыке медицины и информатики.

Еще один проект, который я начала в прошлом году: определение вовлеченности студентов в учебный процесс с помощью нейросетей.

Можно снимать лица студентов на занятии на видеокамеру (данные будут обезличены, с законом об персональных данных всё будет в порядке), затем «прогонять» видео через нейронную сеть, которая будет считывать их мимику, движения и другие параметры и определять, насколько студенты вовлечены. Здесь точность уже будет ниже, но всё же довольно высокая – 80%. Вовлеченность можно будет посчитать в каждый момент занятия, и

преподаватель, таким образом, сможет понять, какие места в его объяснениях были выслушаны внимательно (скажем, с вовлеченностью 95%), а какие – нет (с вовлеченностью 50%), и сделать для себя соответствующие выводы относительно методики своей подачи материала.

- Как правило, анализ вовлеченности делается с помощью анкетирования и уже после занятий. Что натолкнуло тебя на такую идею? Вокруг тебя на занятии сидели не слишком вовлеченные товарищи?

- Заместитель директора бизнес-школы НИЯУ МИФИ Павел Иванович Колыхалов предложил мне эту идею. А, кроме того, моя мама – школьный учитель, и она тоже сказала мне, что хорошо бы иметь такой инструмент контроля вовлеченности учащихся. Анкетирование – это не анализ, что называется, «в моменте», а с помощью разрабатываемого мной инструмента можно сделать достаточно точные измерения в каждый момент объяснения.

- Сделай тогда уж и приложение для анализа вовлеченности преподавателей, когда они принимают зачеты или экзамены, а то нечестно будет...

- Может быть, сделаю. Но серьезно – с этим проектом по вовлеченности я выступала на Весенней научной сессии МИФИ и тоже стала победителем.



- Чем ты думаешь дальше заниматься?

- Я пытаюсь попробовать себя в разных направлениях, занимаюсь, например, дизайном сайтов, разрабатываю цифровые шаблоны для уроков математики. Собственно, моя мама – учитель математики – и направила меня в наш лицей № 1511. В 10 классе я училась в IT-школе Samsung – там мы делали приложения для Android, и моё как раз было связано с онкологией, по шести видам болезней: два варианта онкологии легких (по КТ и по биопсии), онкология полости рта, мозга (по МРТ), молочной железы (по биопсии) и кожи.

- Как выглядят снимки КТ все примерно представляют. А как создается с их помощью такое приложение – алгоритм действий?

- Загружается огромное количество снимков КТ. Затем сверточная нейронная сеть, обучение которой происходит путем обнаружения разнообразных абстрактных характеристик, варьирующихся от простых до сложных, использует эти признаки для прогнозирования правильной категории результата. Основными функциональными особенностями распознавания образов являются нормализация, свертка, активация, объединение, полносвязные слои и классификаторы.

- Откуда ты брала снимки?

- Из открытого банка данных *Kaggle*. А по проекту вовлеченности студентов – находила изображения в интернете сама, размечала их вручную, вместе с моими одноклассниками мы обработали так 845 фотографий студентов на занятиях. Это вышло очень затратно по времени, но зато качественно, мы отбирали фотографии таким образом, чтобы был соблюден баланс классов.

- Но у вас же распознавание по видео?

- Оно же и по фото, так как видео будет разбиваться на кадры, то есть отдельные снимки, на них будут размечаться лица, и прогоняться через нейронную сеть. Абсолютную вовлеченность «студента Васи» мы гарантировать не можем – «Вася», наверное, сам даже не скажет, весь он погружен в процесс или одна мысль об обеде ему всё-таки как-то мешает. Поэтому 100% вовлеченности быть не может, так же как и 0% – даже если студент сидит в наушниках, и дремлет, он хоть иногда видит, что на доске написано. Сейчас дорабатываю этот проект и буду тестировать в этом году на занятиях в нашем университете. Это первый опыт, но он может масштабироваться и быть примененным для бизнеса – можно, например, посчитать процент вовлеченности в рабочие процессы сотрудников какой-то компании. Вот вам и бизнес-информатика. Возможно, из этого вырастет и тема для диплома.

- Что ты любишь вообще делать – для души?

- Выкладывать алмазную мозаику – это как вышивка, только гораздо красивее, и это кропотливая очень работа.

- Теперь понятно, откуда берется терпение, чтобы 800 фотографий вручную обработать. Успехов!

Источник: официальный сайт НИЯУ МИФИ, 14 октября 2024 года

<https://mephi.ru/press/news/23544>