Мини исследование на тему

«Проблемы и перспективы Обучения представлений»

Ерофеев И.А.

Основная идея обучения представления в распутывании клубка факторов на пространстве данных в удобные для обработки гиперплоскости при помощи пространственных преобразований. Для этого нужны некоторые подсказки, априоры, такие как предположения о многообразии представлений, естественной кластеризации, временной и пространственной согласованности и т.п.

Проблемы:

Одна из проблем обучения представления, которая отличает его от других задач машинного обучения, таких как классификация, заключается в трудности установления четкой цели или цели обучения.

У современных NN есть тенденция ловить не высокоуровневые абстракции, а некоторые поверхностные подсказки, не те, которые люди считают наиболее важными. Они эксплуатируют слишком много закономерностей низкого уровня. Как следствие малым зашумлением можно совершенно сбить с толку. Например умные машины могут неправильно распознавать объекты.  
Так же проблема в недостатке обучаемости, или недостатке данных. Так же статистически данные с которыми работает сеть отличаются от тех, на которых обучается. Редкие данные будут выучены хуже. Например кадры около аварийных ситуаций. Но человеку не надо, иногда даже бывать в таких ситуациях, или видеть их, что бы знать, что в них делать. Мы можем обобщать интуитивно.   
Мы обладаем более абстрактным мышлением, и глубоким пониманием.

Проблема обучения без учителя так же в том что обучение идёт в области данных, а не в абстрактной области, т.к. используется в качестве функции потерь ошибка реконструкции, или ошибка «похожести». Если перевести маппинг из области данных в абстрактную область, это может дать возможность строить более удобные или правильные планы. Пример такой проблемы – при чистом обучении без учителя на речи, сеть ловит текстуру речи но не ловит фонемы т.к. текстура передаётся порядка 16 тысяч чисел в секунду простив нескольких фонем в секунду. Решением будет обучить с учителем преобразование фонем в акустику и отдельно на порядок фонем во фразе.

Перспективы:  
Один из способов найти некоторые из факторов хорошего представления – подсказки о существовании этих факторов через управление ими. Мы можем воздействовать на конкретные аспекты миры, (не сильно влияя при этом на остальной мир). Например позиция объекта в пространстве. Но держать по нейрону для каждого объекта это расточительно, потому важно обобщать, т.е. переходить от факторов к некоторым именам. Запоминать понятие «позиция» а не конкретные позиции каждого объекта. После чего использовать уже вектора понятий как эмбединги.

Агентное обучение:  
Дети взаимодействуют с миром, а не просто наблюдают. Таким образом они например создают довольно точную модель физики. и представлять это после. Планировать, опираясь на эти представления. Мы действуем. У нас есть знание об агенте, и это субъективное знание, зависящее от самого агента. Например кто-то может управлять автомобилем, а кто-то нет. Кто-то лучше, кто-то хуже. И это наше отличие от AI, который опирается только на объективное знание.

То что человек представляет имеет много размерностей, но состоит из небольшого кол-ва абстрактных понятий о мире.  
Таким образом можно высказать предположение что много может быть сказано о мире в одном малоразмерном предложении (обращаясь всего к нескольким переменным)   
Например если уронить ручку со стола, мы можем не знать куда конкретно она приземлиться, но мы можем предсказать что она упадёт на пол.  
  
Пример организации перспективной памяти: маломерная сознательная память, которая хранит абстракции и через некоторый механизм внимания активизирует многомерную бессознательную память. И обучается соответственно механизм обновления памяти и внимания.

Ближайшие перспективы: Использование GAN, а так же анализ взаимной информации.   
Современная проблема GAN: тенденция создавать хорошие объекты но терять некоторые области. Например генератор научиться генерировать только 0 и 8 из всех цифр. Одно из решений этой проблемы – использовать взаимную информацию в качестве регуляризатора генератора (максимизация энтропии выхода).   
Идея: не генерировать изображения, а делать представления и скрытые представления которые не обязательно ловят всё, но ловят самое главное, описывающее большую часть информации.  
Если провести прямую в области данных между точками различных классов, то скорее всего большая часть точек не будут представителями ни одного из классов. В то время как если взять правильное представление, то большинство точек на такой прямой будут принадлежать какому-то классу.

Ещё одна идея: самообучение (Self-supervise)   
Любой сигнал. Предсказание какой-то его части по другим частям.  
Например порядок картинок в серии или с этой ли картинки этот кусочек, и куда его вставить и т.п.  
  
Таким образом задавать вопросы не в области данных а в некоторой латентной области.  
  
Соответственно из акустического сигнала предсказывать следующий, или из фонемы след фонему (различные уровни).

Так же можно использовать данные подход не только на уровне сегментов и кусочков, а на множестве уровней, и таким образом улавливать признаки принадлежащие множеству уровней.

Так же можно создавать глобальный ввектор признаков + вектора кусочков и соответственно сравнивать не только вектора кусочков между собой, но и с глобальным вектором. (Как например глобальный признак автомобиль, а на локальном кусочке только, например, скамейка в далеке).

Возникает сложный вопрос: Как понять что всё работает хорошо?

Примеры решений:  
1) если эти признаки (как глобальные так и локальные) подавать на классификатор для предсказания важного фактора.

(Получившийся таким образом классификатор иной раз работает даже лучше, чем обучение с учителем.)

2) измерять взаимную информацию и зависимость

3) Реконструировать по признакам объект.

Основная проблематика последних лет: сделать обучение без учителя сравнимым или лучше чем с учителем.

На данный момент обучение с учителем для конкретной задачи всегда будет лучше т.к. метка, поставленная человеком не случайная функция данных, а конкретная, важная функция, содержащая высокоуровневую концепцию, важную для человека.   
  
Однако у такого обучения есть проблемы, и несколько главных:  
1) Как только найден «неплохой» признак для классификации – обучение останавливается (локальный минимум). Для конкретной задачи (конкретных фонем например) это хорошо.  
Однако при добавлении новых фонем в список, этот признак может оказаться непригоден, в то время как обучение без учителя даёт более богатый набор признаков.

2) Переобучение. В то время как в обучении без учителя неизвестно что конкретно важно => потенциально более стабильно для связанных задач, не используемых в тренировке.

Происходит обобщение не внутри распределения, а через связанные распределения.

Современные ограничения на примере языко-ориентированных задач:

Существует 2 условных типа мышления: Быстрое, интуитивное, не объяснимое словами и медленное, логичное, последовательное, объяснимое. Распознование лиц и математические расчёты. Классические алгоритмы занимаются вторым, глубокое обучение первым. Вопрос в синтезе. Для понимания естественных языков нужно совместное развитие всех аспектов AI.   
Один из аспектов, не так часто встречающихся, который тем не менее упоминается выше, это то, как агенты понимают мир через поимку причинно-следственных структур, то чего пока не умеет ML. «Что будет если я или он сделает то-то?»  
  
Наиболее интересные источники: Соответствующие топики на <https://www.quora.com/>  
а так же <https://www.microsoft.com/en-us/research/blog/>  
Важным было разобраться в целом что это такое, и как это работает, с чем очень помог первый ресурс.