

TensorFlow Overview

Назначение библиотеки TensorFlow

Библиотека TensorFlow предназначена для использования в решения задач в области машинного обучения. Библиотека предоставляет инструменты для широкого семейства техник машинного обучения.

Таблица плюсов и минусов.

Параметр	Оценка (+/-)
Использование GPU для вычислений	+
Внятная документация	+
В процессе обучения есть возможность сохранять состояние модели (checkpoints)	+
Встроенная система визуализации вычислительного процесса и результатов (TensorBoard)	+
Удобная система подачи данных (датасеты)	+
Поддержка OpenCL «из коробки»	-
При использовании сторонних библиотек (при чтении использовался openCV) наблюдалось сильное падение производительности	-
Отсутствие инструментов для создания сверточных слоев CNN для изображений с гексагональными пикселями	-

Дополнительно:

Типы и механизмы представления входных данных

В библиотеке представлены 9 классов. Основные их виды:

- числовые
- колонки индикаторов/категорий (one-hot vector)
- вложенные (embedding column) - позволяют понизить размерность используя таблицу поиска (lookup table),
- хеширование для большого количества категорий/сетки
- Bucketized column — для ввода сетки

Датасеты

Удобная система для подготовки данных для чтения. Имеет возможность нативно реализовать чтение как из отдельных файлов так и из множества файлов.

Создание собственного эстиматора

Создание собственного эстиматора требует создания всего одной функции `model_fn`.

Основные части функции `model_fn`:

- определение входного слоя
- определение модели
- определение выходного слоя

- предсказание
- определение функции потерь
- определение метрик(и) оценки качества
- выбор алгоритма минимизации

Создание эстиматора облегчает работу с библиотекой.