

Кастомная категория биржевых акций

Сервис ML для фильтрации акций по
произвольному пользовательскому принципу

Категории биржевых акций

- Без ML пользователь получает фиксированный список категорий от конкретного брокера, например: ресурсы, производство, медицина, финансы и т.д.. Но у каждого инвестора подход отличается, поэтому может потребоваться своя кастомная категория под конкретные цели.
- С ML сервисом пользователь может обучить модель распознавать логику кастомной категории акций (по любой пользовательской логике, на примерах) и получить свою кастомную категорию акций для целей выбранного подхода инвестирования.

Требования и ограничения

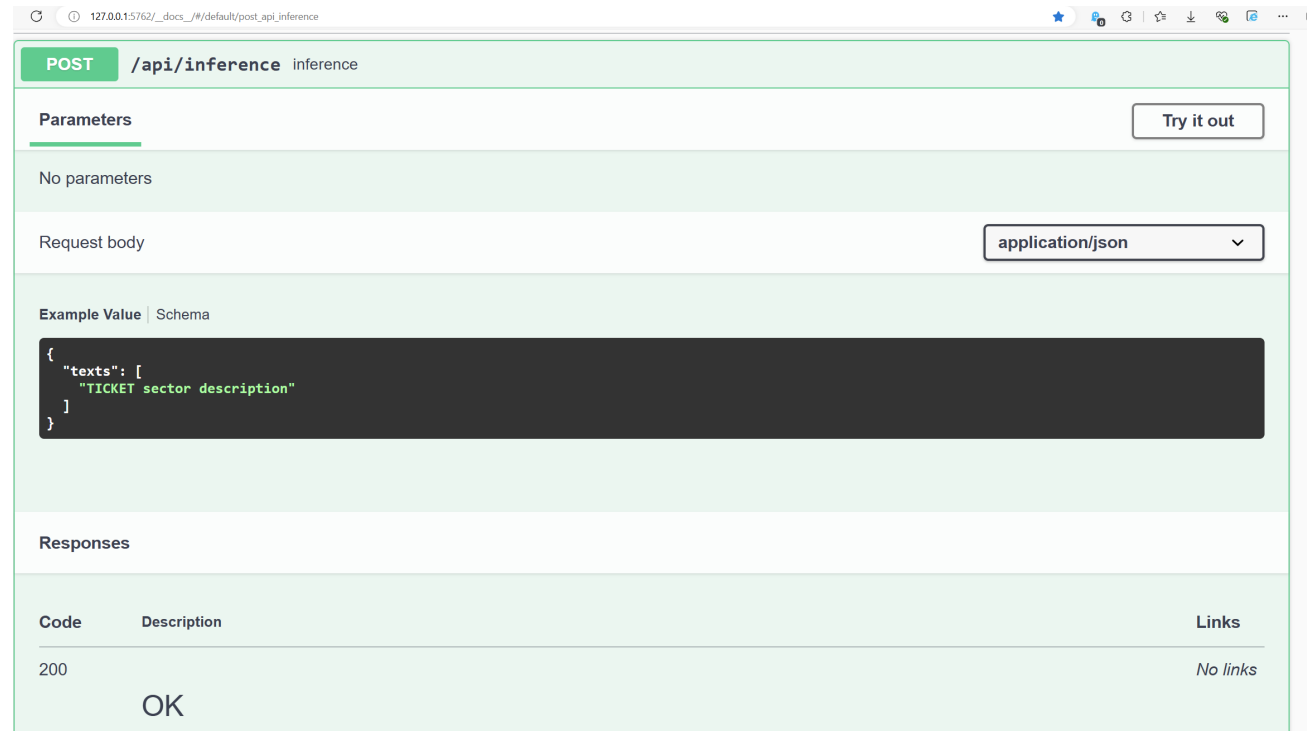
- Важен recall, а не precision: инвестор хочет найти все нужные акции, а “лишние” акции инвестициям не мешают.
- Модель выдает позитивный ответ на пустую запись или на запись с только неизвестными модели словами. Это не является недостатком, потому что приоритет на sensitivity, а такая стратегия дает покрытие 100 %.

Обучение модели

- Обучение модели происходит на примерах, пользователю достаточно поставить метки у некоторого набора примеров: 1 — выбираем акцию; 0 — не выбираем акцию.
- Модель анализирует текст описаний компаний + категорию от брокера. Пример текста описания (в начале описания название категории от брокера):
„industrials ПАО «Яковлев» занимается производством гражданских и военные летательных аппаратов, в том числе многоцелевых боевых самолётов Су-30МК,...”

ML микросервис

- На выходе пользователь получает REST JSON микросервис, который можно развернуть в среде с поддержкой Docker .



Достижение целей

- Пользователь может применять микросервис отдельно или интегрировать в развитую программную систему.
- Пользователь получает возможность отбирать акции по уникальной кастомной категории с целью инвестиций по выбранной стратегии.

Также, ML модель позволяет:

- расширить число акций в активе, потому что модель фильтрует акции беспристрастно и внимательно с высоким покрытием, исключается человеческий фактор;
- пользователь автоматически получит информацию о новых появляющихся на бирже компаниях, которые будут захвачены классификатором;
- исследовать рынок по выбранному категориальному критерию, потому что модель может «заметить» акции, на которые не обратил внимание человек (например, в случаях, когда неочевидно, что акция на самом деле подходит под критерии, а модель такие случаи может обнаружить) и привести к некоторым «открытиям»;
- также, модель экономит время при многократном/периодическом использовании inference, потому что человеку пришлось бы неоднократно просматривать и анализировать информацию по всем акциям, на такую работу ушло бы больше времени.