

Пятая конференция разработчиков ПО «DevParty»

25 марта 2017 года, Вологда



DEV PARTY
конференция разработчиков

Статистические предсказания игровой активности

Иванов Игорь

Playrix

Аналитика

- Апостериорный анализ:
 - Эксперимент
 - Оценка результата
- Априорный анализ:
 - Моделирование эксперимента
 - Прогнозирование результата

Какую информацию “из
будущего” по игрокам
хотелось бы получить?

- Увлечёт ли игра?
- Что понравится больше, а что меньше?
- Можно ли ожидать конверсию?
- С какой скоростью игрок будет проходить игру?
- Когда игрок перестанет играть?
- Хватит ли терпения на прохождение сложных уровней?
- ... (тысячи их)

Где найти
хрустальный шар?

Статистические модели предсказаний

- на известных данных определяем зависимости (**обучение**) и применяем их к неизвестным данным (**предсказание**)

- Выявим паттерны поведения
- Определим критерии “похожести” игроков
- Определим критерии разделения полезной информации и шума
- По полезной информации учимся строить прогнозы
- ...
- Profit!

Фундамент для предсказательных моделей

- Теория вероятностей (изучаем случайное явление само по себе)
- Математическая статистика (изучаем предыдущие эксперименты в которых наблюдалось случайное явление)
- Факторный анализ (изучаем влияние других явлений на случайное явление, чтобы объяснить предыдущие эксперименты)

- Регрессия - предсказываем количественный признак
- Классификация - предсказываем качественный признак
- Кластеризация - определяем принадлежность ко множеству

Интерпретируемые модели

- Определяем что можно изменить чтобы достичь нужного результата
- Отсекаем значимое от второстепенного
- Уменьшаем вычислительные затраты
- Увеличиваем устойчивость

Менее интерпретируемые модели

- Повышаем точность предсказания
- Упрощаем отбор значимых данных
- Повышаем универсальность решения
- Увеличиваем размерность данных

Как измерить точность
предсказания?

- Метод минимизации эмпирического риска (“хорошо - это когда мало ошибок”)
- Метод максимального правдоподобия (“хорошо - это когда наиболее часто встречается наиболее вероятное значение”)
- Минимизация ошибок определённого рода

А какая точность нас
устроит?

Предсказание: все
присутствующие согласятся с
тем, что нельзя предсказывать со
100% точностью :)

- Интуиции доверять не стоит
- Как бы много мы не знали про игрока, остаются неизвестные факторы
- Глубокое знание предметной области помогает выбрать правильные критерии оценки

Почему нельзя
предсказывать всё
подряд?

Принцип Бонферрони

- Вычислить количество рассматриваемых событий при предположении их полной случайности
- Если полученное количество намного превосходит количество реальных искомых событий, то найденные события фиктивные

- Нет смысла в информации, которую мы не знаем как использовать
- Нет смысла в предсказании, которое мы не сможем воспроизвести
- Нет смысла в очевидном предсказании

Примеры

- Предсказание “неконверсии”
- Предсказание отвалов игроков с большим количеством покупок
- Предсказание отвалов на paywall

Спасибо!