

# PRESENCE-ONLY DATA И ДРУГИЕ ПРИМЕРЫ ЕМ

---

Сергей Николенко

СПбГУ — Санкт-Петербург

16 ноября 2021 г.

---

*Random facts:*

- 16 ноября в ООН — Международный день толерантности, который отмечается и в России; например, в 2011 году в конкурсе «Как быть толерантным», проводившемся в 420 школе Санкт-Петербурга, победила команда восьмиклассников под названием «14/88»; почётная грамота долго висела в фойе школы
- 16 ноября 1805 г. в рамках Войны третьей коалиции произошёл Шёнграбенский бой, во многом сделавший имя Багратиону
- 16 ноября 1871 в штате Нью-Йорк редактор Army and Navy Journal Уильям Чёрч и капитан Джордж Уингейт официально зарегистрировали National Rifle Association; сейчас в NRA состоит более 5 миллионов человек, и это одна из самых могущественных лоббистских групп Вашингтона
- 16 ноября 1938 г. — день рождения самбо; в этот день Спорткомитет включил самбо в число видов спорта, культивируемых в СССР
- 16 ноября 1974 г. Александр Солженицын провёл в Цюрихе пресс-конференцию «О будущем России», предложив «Программу "нравственной революции"», изложенную в эссе «Жить не по лжи»

## ПРИМЕР EM: PRESENCE-ONLY DATA

---

## PRESENCE-ONLY DATA

- Пример из экологии: пусть мы хотим оценить, где водятся те или иные животные.
- Как определить, что суслики тут водятся, понятно: видишь суслика — значит, он есть.
- Но как определить, что суслика нет? Может быть, ты не видишь суслика, и я не вижу, а он есть?..



## PRESENCE-ONLY DATA

- Формально говоря, есть переменные  $\mathbf{x}$ , определяющие некий регион (квадрат на карте), и мы моделируем вероятность того, что нужный вид тут есть,  $p(y = 1 | \mathbf{x})$ , при помощи логит-функции:

$$p(y = 1 | \mathbf{x}) = \sigma(\eta(\mathbf{x})) = \frac{1}{1 + e^{-\eta(\mathbf{x})}},$$

где  $\eta(\mathbf{x})$  может быть линейной (тогда получится логистическая регрессия), но может, в принципе, и не быть.

- Заметим, что даже если бы мы знали настоящие  $y$ , это было бы ещё не всё: сэмплирование положительных и отрицательных примеров неравномерно, перекошено в пользу положительных.
- *Важное замечание: prospective vs. retrospective studies, case-control studies.*

- Это значит, что ещё есть пропорции сэмплирования (sampling rates)

$$\gamma_0 = p(s = 1 | y = 0), \quad \gamma_1 = p(s = 1 | y = 1),$$

т.е. вероятности взять в выборку положительный/отрицательный пример.

- Их можно оценить как

$$\gamma_0 = \frac{n_0}{(1 - \pi)N}, \quad \gamma_1 = \frac{n_1}{\pi N},$$

где  $\pi$  — истинная доля положительных примеров (встречаемость, occurrence).

- Кстати, эту  $\pi$  было бы очень неплохо оценить в итоге.

- Тогда, если знать истинные значения всех  $y$ , то в принципе можно обучить:

$$\begin{aligned}
 p(y = 1 \mid s = 1, \mathbf{x}) &= \\
 &= \frac{p(s = 1 \mid y = 1, \mathbf{x})p(y = 1 \mid \mathbf{x})}{p(s = 1 \mid y = 0, \mathbf{x})p(y = 0 \mid \mathbf{x}) + p(s = 1 \mid y = 1, \mathbf{x})p(y = 1 \mid \mathbf{x})} = \\
 &= \frac{\gamma_1 e^{\eta(\mathbf{x})}}{\gamma_0 + \gamma_1 e^{\eta(\mathbf{x})}} = \frac{e^{\eta^*(\mathbf{x})}}{1 + e^{\eta^*(\mathbf{x})}},
 \end{aligned}$$

где  $\eta^*(\mathbf{x}) = \eta(\mathbf{x}) + \log(\gamma_1/\gamma_0)$ , т.е.

$$\eta^*(\mathbf{x}) = \eta(\mathbf{x}) + \log\left(\frac{n_1}{n_0}\right) - \log\left(\frac{\pi}{1 - \pi}\right).$$

- Таким образом, если  $\pi$  неизвестно, то  $\eta(\mathbf{x})$  можно найти с точностью до константы.
- А у нас не  $n_0$  и  $n_1$ , а naive presence  $n_p$  и background  $n_u$ , т.е.

$$p(y = 1 | s = 1) = \frac{n_p + \pi n_u}{n_p + n_u}, \quad p(y = 1 | s = 0) = \frac{(1 - \pi)n_u}{n_p + n_u},$$

$$\gamma_1 = \frac{p(y=1|s=1)p(s=1)}{p(y=1)} = \frac{n_p + \pi n_u}{\pi(n_p + n_u)}p(s = 1),$$

$$\gamma_0 = \frac{p(y=0|s=1)p(s=1)}{p(y=0)} = \frac{n_u}{n_p + n_u}p(s = 1),$$

и в нашей модели  $\log \frac{n_1}{n_0} = \log \frac{n_p + \pi n_u}{\pi n_u}$ , т.е. всё как раньше, но  $n_1 = n_p + \pi n_u$ ,  $n_0 = (1 - \pi)n_u$ .

- Обучать по  $y$  можно так: обучить модель, а потом вычесть из  $\eta(\mathbf{x})$  константу  $\log \frac{n_p + \pi n_u}{\pi n_u}$ .

- Но у нас нет настоящих данных  $y$ , чтобы обучить регрессию, а есть только presence-only  $z$ : если  $z = 1$ , то  $y = 1$ , но если  $z = 0$ , то неизвестно, чему равен  $y$ .
- (Ward et al., 2009): давайте использовать EM. Правдоподобие:

$$\begin{aligned}\mathcal{L}(\eta \mid \mathbf{y}, \mathbf{z}, X) &= \prod_i p(y_i, z_i \mid s_i = 1, \mathbf{x}_i) = \\ &= \prod_i p(y_i \mid s_i = 1, \mathbf{x}_i) p(z_i \mid y_i, s_i = 1, \mathbf{x}_i).\end{aligned}$$

- В нашем случае при  $n_p$  положительных примеров и  $n_u$  фоновых (неизвестных)

$$\begin{aligned}p(z_i = 0 \mid y_i = 0, s_i = 1, \mathbf{x}_i) &= 1, \\ p(z_i = 1 \mid y_i = 1, s_i = 1, \mathbf{x}_i) &= \frac{n_p}{n_p + \pi n_u}, \\ p(z_i = 0 \mid y_i = 1, s_i = 1, \mathbf{x}_i) &= \frac{\pi n_u}{n_p + \pi n_u}.\end{aligned}$$

- А максимизировать нам надо сложное правдоподобие, в котором значения  $\mathbf{y}$  неизвестны:

$$L(\eta | \mathbf{z}, X) = \prod_i p(z_i | s_i = 1, \mathbf{x}) =$$

$$= \prod_i \left( \frac{\frac{n_p}{\pi n_u} e^{\eta(\mathbf{x}_i)}}{1 + \left(1 + \frac{n_p}{\pi n_u}\right) e^{\eta(\mathbf{x}_i)}} \right)^{z_i} \left( \frac{1 + e^{\eta(\mathbf{x}_i)}}{1 + \left(1 + \frac{n_p}{\pi n_u}\right) e^{\eta(\mathbf{x}_i)}} \right)^{1-z_i} .$$

- Для этого и нужен EM.

- E-шаг здесь в том, чтобы заменить  $y_i$  на его оценку

$$\hat{y}_i^{(k)} = \mathbb{E} [y_i \mid \eta^{(k)}] = \frac{e^{\eta^{(k)}} + 1}{1 + e^{\eta^{(k)}} + 1}.$$

- M-шаг мы уже видели, это обучение параметров логистической модели с целевой переменной  $\mathbf{y}^{(k)}$  на данных  $X$ .

(1) Chose initial estimates:  $\hat{y}_i^{(0)} = \pi$  for  $z_i = 0$ .

(2) Repeat until convergence:

- *Maximization step:*

– Calculate  $\hat{\eta}^{(k)}$  by fitting a logistic model of  $\hat{\mathbf{y}}^{(k-1)}$  given  $X$ .

– Calculate  $\hat{\eta}^{(k)} = \hat{\eta}^{*(k)} - \log\left(\frac{n_p + \pi n_u}{\pi n_u}\right)$ .

- *Expectation step:*

$$\hat{y}_i^{(k)} = \frac{e^{\hat{\eta}^{(k)}}}{1 + e^{\hat{\eta}^{(k)}}} \text{ for } z_i = 0 \quad \text{and} \quad \hat{y}_i^{(k)} = 1 \text{ for } z_i = 1$$

- У Ward et al. получалось хорошо, но тут вышел любопытный спор.
- Ward et al. писали так: хотелось бы, чтобы можно было оценить  $\pi$ , но “ $\pi$  is identifiable only if we make unrealistic assumptions about the structure of  $\eta(\mathbf{x})$  such as in logistic regression where  $\eta(\mathbf{x})$  is linear in  $\mathbf{x}$ :  $\eta(\mathbf{x}) = \mathbf{x}^T \beta$ ”.
- Через пару лет вышла статья Royle et al. (2012), тоже очень цитируемая, в которой говорилось: “logistic regression... is hardly unrealistic... such models are the most common approach to modeling binary variables in ecology (and probably all of statistics)... the logistic functions... is customarily adopted and widely used, and even books have been written about it”.

- Они предложили процедуру для оценки встречаемости  $\pi$ :
  - для признаков  $\mathbf{x}$  у нас  $p(y = 1 | \mathbf{x}) = \frac{p(y=1)\pi_1(\mathbf{x})}{p(y=1)\pi_1(\mathbf{x})+(1-p(y=1))\pi_0(\mathbf{x})}$ ;
  - данные – это выборка из  $\pi_1(\mathbf{x})$  и отдельно выборка из  $\pi(\mathbf{x})$ ;
  - как видно, даже если знать  $\pi$  и  $\pi_1$  полностью, остаётся свобода: надо оценить  $p(y = 1 | \mathbf{x})$ , а  $\pi_0(\mathbf{x})$  мы не знаем;
  - Royle et al. вводят предположения для  $p(y = 1 | \mathbf{x})$  в виде логистической регрессии:  $p(y = 1 | \mathbf{x}) = \sigma(\beta^\top \mathbf{x})$ ;
  - тогда действительно можно записать  $p(y = 1)\pi_1(\mathbf{x}) = p(y = 1 | \mathbf{x})\pi(\mathbf{x}) = p(y = 1, \mathbf{x})$ , и если  $\pi(\mathbf{x})$  равномерно (это логично), то

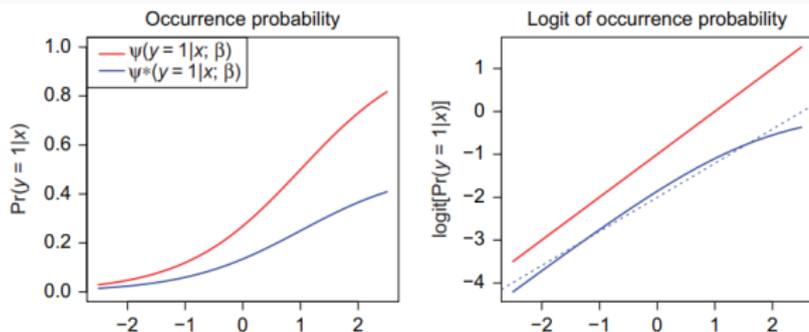
$$\pi_1(\mathbf{x}_i) = \frac{p(y_i = 1 | \mathbf{x}_i)}{\sum_{\mathbf{x}} p(y = 1 | \mathbf{x})};$$

если подставить сюда логистическую регрессию, то можно оценить  $\beta$  максимального правдоподобия, и не надо знать  $p(y = 1)$ !

- Что тут не так?

## PRESENCE-ONLY DATA

- Всё так, но предположение о логистической регрессии здесь выполняет слишком много работы.
- Если рассмотреть две кривые, у которых  $p^*(y = 1 | \mathbf{x}, \beta) = \frac{1}{2}p(y = 1 | \mathbf{x}, \beta)$ , то у них будет  $p^*(y = 1) = \frac{1}{2}p(y = 1)$ , но общее правдоподобие  $\pi_1(\mathbf{x}_i)$  будет в точности одинаковое,  $\frac{1}{2}$  сократится.
- Дело в том, что модель  $p^*$  не будет логистической регрессией, с  $\beta$  произойдёт что-то нелинейное; но откуда у нас настолько сильное предположение? Как отличить синюю кривую справа от пунктирной прямой?



# РЕЙТИНГ-СИСТЕМЫ, ОБУЧАЮЩИЕСЯ ММ-АЛГОРИТМАМИ

---

- Рейтинг-система — это модель, которая ранжирует участников (игроков) в единый линейный порядок по данным сравнений небольших подмножеств этих игроков (турниров).
- Более того, результаты турниров зашумлены (отчасти случайны).
- Соответственно, и применяются они в таких ситуациях (пример: контекстная реклама в Bing).

- Первая известная рейтинг-система, основанная на байесовском подходе — это рейтинг Эло.
- Суть модели:
  - сила игры шахматиста в одной партии — случайная величина;
  - *рейтинг* — это ожидание этой величины; мы пытаемся оценить это ожидание;
  - исходная модель Эло — нормальное распределение силы игры вокруг рейтинга.

- Значит, сила игры в конкретной партии распределена как

$$p(x) = N(x; s, \beta) = \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{1}{2\beta^2}(x-s)^2}.$$

- Сила игры задаётся двумя параметрами: средним  $s$  (собственно рейтингом) и дисперсией  $\beta^2$ .
- Эло предположил, что дисперсия  $\beta^2$  постоянна (и даже от игрока не зависит), а среднее — это как раз рейтинг, который мы пытаемся оценить.

- Значит, математически говоря, мы ищем

$$\begin{aligned}\arg \max_{s, \beta^2} p(s, \beta^2 | D) &= \arg \max_{s, \beta^2} \frac{p(D | s, \beta^2) p(s, \beta^2)}{p(D)} = \\ &= \arg \max_{s, \beta^2} p(D | s, \beta^2) p(s, \beta^2).\end{aligned}$$

- Как мы знаем, нормальное распределение является самосопряжённым, поэтому если сила игры нормально распределена вокруг рейтинга, то логично взять нормальное распределение как априорное для рейтинга.
- Таким образом, рейтинг игрока складывается из двух чисел: его среднего значения  $\mu$  и дисперсии  $\sigma^2$ .
- Значение  $\mu$  отображается в таблице рейтингов, а  $\sigma^2$  показывает, насколько достоверна имеющаяся оценка.

- Предположим, что встречаются два игрока с некоторыми априорными распределениями на рейтинги  $N(s_1; \mu_1, \sigma_1^2)$  и  $N(s_2; \mu_2, \sigma_2^2)$ .
- Тогда сила игры каждого из них в этой конкретной партии имеет распределение

$$\begin{aligned} p(x | \mu, \sigma) &= \int_{-\infty}^{\infty} p(x | s)p(s | \mu, \sigma)ds = \\ &= \int_{-\infty}^{\infty} N(x; s, \beta)N(s; \mu, \sigma)ds = \\ &= \int_{-\infty}^{\infty} \frac{1}{\beta\sqrt{2\pi}}e^{-\frac{1}{2\beta^2}(x-s)^2} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(s-\mu)^2} ds = N(x; \mu_x, \sigma_x). \end{aligned}$$

то есть мы снова приходим к нормальному распределению, но с другими параметрами.

- Задача обучения заключается в том, чтобы после новой партии принять во внимание её результат и пересчитать рейтинги.
- Эло разработал специальные аппроксимации и очень простые алгоритмы для этого случая (через «ожидаемые очки в турнире»), чтобы каждый шахматист мог сам на калькуляторе свой рейтинг посчитать, но они нас сейчас не очень интересуют.
- Сейчас есть обобщения рейтинга Эло, мы поговорим о них позже.

- Другой подход к рейтинг-системам — модели Брэдли-Терри (Bradley-Terry).
- Модель предполагает, что для участников  $1, \dots, n$  можно подобрать такие рейтинги  $\gamma_i, i = 1..n$ , что вероятность победы участника  $i$  над участником  $j$  равна

$$p(i \text{ побеждает } j) = \frac{\gamma_i}{\gamma_i + \gamma_j}.$$

- Основная задача заключается в том, чтобы найти  $\gamma = (\gamma_1, \dots, \gamma_m)$  максимального правдоподобия из имеющихся данных  $D$ .

- Если принять априорное распределение равномерным, можно просто максимизировать правдоподобие

$$p(D|\gamma) = \prod_{i=1}^m \prod_{j=1}^m \left( \frac{\gamma_i}{\gamma_i + \gamma_j} \right)^{w_{ij}},$$

где  $w_{ij}$  — то, сколько раз  $x_i$  обыграл  $x_j$  при их попарном сравнении ( $w_{ii} = 0$  по определению).

- Взяв логарифм, будем максимизировать

$$l(\gamma) = \sum_{i=1}^m \sum_{j=1}^m (w_{ij} \log \gamma_i - w_{ij} \log(\gamma_i + \gamma_j)).$$

- Максимизировать будем классическим ММ-алгоритмом (minorization-maximization), фактически вариационным приближением.
- Заметим, что

$$1 + \log \frac{x}{y} - \frac{x}{y} \leq 0.$$

**Упражнение.** Докажите это.

- Рассмотрим вспомогательную функцию

$$Q(\gamma, \gamma^{(k)}) = \sum_{i,j} w_{ij} \left[ \log \gamma_i - \frac{\gamma_i + \gamma_j}{\gamma_i^{(k)} + \gamma_j^{(k)}} - \log (\gamma_i^{(k)} + \gamma_j^{(k)}) + 1 \right].$$

**Упражнение.** Используя предыдущее неравенство, докажите, что  $Q(\gamma, \gamma^{(k)}) \leq l(\gamma)$ .

- Чтобы найти  $\max_{\gamma} Q(\gamma, \gamma^{(k)})$ , можно просто взять производные  $\frac{\partial Q}{\partial \gamma_l}$ :

$$\begin{aligned} \frac{\partial Q}{\partial \gamma_l} &= \sum_{i,j} w_{ij} \left[ \frac{\delta_{il}}{\gamma_i} - \frac{\delta_{il} + \delta_{jl}}{\gamma_i^{(k)} + \gamma_j^{(k)}} \right] = \\ &= \frac{1}{\gamma_l} \sum_j w_{lj} - \sum_j \frac{w_{lj}}{\gamma_i^{(k)} + \gamma_j^{(k)}} - \sum_i \frac{w_{il}}{\gamma_i^{(k)} + \gamma_j^{(k)}}. \end{aligned}$$

- Если  $w_l$  — общее количество побед игрока  $l$  ( $w_l = \sum_j w_{lj}$ ), и  $N_{ij}$  — количество встреч между игроками  $i$  и  $j$  ( $N_{ij} = w_{ij} + w_{ji}$ ), получаем

$$\frac{w_l}{\gamma_l} - \sum_j \frac{N_{lj}}{\gamma_l^{(k)} + \gamma_j^{(k)}} = 0.$$

- В результате правило пересчёта на одной итерации выглядит так:

$$\gamma_l^{(k+1)} := w_l \left[ \sum_j \frac{N_{lj}}{\gamma_l^{(k)} + \gamma_j^{(k)}} \right]^{-1}.$$

- Получили алгоритм оценки рейтингов. Правда, он пока работает только для ситуации, когда игроки встречаются один на один и выигрывают или проигрывают.
- Но даже в шахматах бывают ничьи. Как их учесть?

- Если возможны ничьи, их вероятность можно описать дополнительным параметром  $\theta > 1$ , и это приводит к модели, в которой

$$p(i \text{ побеждает } j) = \frac{\gamma_i}{\gamma_i + \theta\gamma_j},$$

$$p(j \text{ побеждает } i) = \frac{\gamma_j}{\theta\gamma_i + \gamma_j},$$

$$p(i \text{ и } j \text{ играют вничью}) = \frac{(\theta^2 - 1)\gamma_i\gamma_j}{(\gamma_i + \theta\gamma_j)(\theta\gamma_i + \gamma_j)}.$$

- Аналогично можно вводить другие обобщения.
- Например, если результат может зависеть от порядка элементов в паре (скажем, команды проводят «домашние» матчи и «гостевые»), можно ввести дополнительный параметр  $\theta$ , характеризующий, насколько большое преимущество дают «родные стены», и рассмотреть модель с

$$p(i \text{ побеждает } j) = \begin{cases} \frac{\theta\gamma_i}{\theta\gamma_i + \gamma_j}, & \text{если } i \text{ играет дома,} \\ \frac{\gamma_j}{\theta\gamma_i + \gamma_j}, & \text{если } j \text{ играет дома.} \end{cases}$$

- Можно даже обобщить на случай, когда в одном турнире встречаются несколько игроков: пусть перестановка  $\pi$  подмножества игроков  $A = \{1, \dots, k\}$  (результат турнира) имеет вероятность

$$p_A(\pi) = \prod_{i=1}^k \frac{\gamma_{\pi(i)}}{\gamma_{\pi(i)} + \gamma_{\pi(i+1)} + \dots + \gamma_{\pi(k)}}.$$

- Можно показать, что такая модель эквивалентна весьма естественной «аксиоме Люса»: для любой модели, в которой вероятности игроков обыграть друг друга в любой паре не равны нулю, для любых подмножеств игроков  $A \subset B$  и любого игрока  $i \in A$

$$\begin{aligned} p_B(i \text{ побеждает}) &= \\ &= p_A(i \text{ побеждает})p_B(\text{побеждает кто-то из множества } A). \end{aligned}$$

- Но для крупных турниров это перестаёт работать; и совсем трудно что-то осмысленное сделать, если игроки соревнуются не поодиночке, а в командах.

- Пример из практики: в какой-то момент я хотел сделать рейтинг спортивного «Что? Где? Когда?»:
  - участвуют команды по  $\leq 6$  человек, причём часто встречаются неполные команды;
  - игроки постоянно переходят между командами (поэтому TrueSkill);
  - в одном турнире могут участвовать до тысячи команд (синхронные турниры);
  - командам задаётся фиксированное число вопросов (36, 60, 90), т.е. в крупных турнирах очень много команд делят одно и то же место.

- Первое решение — система TrueSkill
- Расскажу о ней потом, когда будем говорить об Expectation Propagation
- У неё есть проблемы в постановке спортивного ЧГК, мы когда-то их отчасти решили, была сложная и интересная модель, она работала лучше базового TrueSkill
- Но...

- В какой-то момент база турниров ЧГК стала собирать повопросные результаты.
- Мы теперь знаем, на какие именно вопросы ответила та или иная команда.
- Так что когда я вернулся к задаче построения рейтинга ЧГК, задача стала существенно проще.

- Пример аналогичного приложения:
  - есть набор вопросов для теста из большого числа вопросов (например, IQ-тест или экзамен по какому-то предмету);
  - участники отвечают на случайное подмножество вопросов;
  - надо оценить участников, но уровень сложности вопросов нельзя заранее точно сбалансировать.
- «Что? Где? Когда?» — это оно и есть, только теперь участники объединяются в команды.

- Baseline — логистическая регрессия:
  - каждый игрок  $i$  моделируется скиллом  $s_i$ ,
  - каждый вопрос  $q$  моделируется сложностью («лёгкостью»)  $c_q$ ,
  - добавим глобальное среднее  $\mu$ ,
  - обучим логистическую модель

$$p(x_{tq} \mid s_i, c_q) \sim \sigma(\mu + s_i + c_q)$$

для каждого игрока  $i \in t$  команды-участницы  $t \in T^{(d)}$  и каждого вопроса  $q \in Q^{(d)}$ , где  $\sigma(x) = 1/(1 + e^x)$  — логистический сигмоид,  $x_{tq}$  — ответила ли команда  $t$  на вопрос  $q$ .

- Логистическая модель предполагает фактически, что каждый игрок ответил на каждый вопрос, который взяла команда.
- Это неправда, мы не знаем, кто ответил, только знаем, что кто-то это сделал; а если команда не ответила, то никто.
- Это по идее похоже на presence-only data models (Ward et al., 2009; Royle et al., 2012).

- Поэтому давайте сделаем модель со скрытыми переменными.
- Для каждой пары игрок-вопрос, добавим переменную  $z_{iq}$ , которая означает, что «игрок  $i$  ответил на вопрос  $q$ ».
- На эти переменные есть такие ограничения:
  - если  $x_{tq} = 0$ , то  $z_{iq} = 0$  для каждого игрока  $i \in t$ ;
  - если  $x_{tq} = 1$ , то  $z_{iq} = 1$  для по крайней мере одного игрока  $i \in t$ .

- Параметры модели те же — скилл и сложность вопросов:

$$p(z_{iq} | s_i, c_q) \sim \sigma(\mu + s_i + c_q).$$

- Обучаем EM-алгоритмом:

- E-шаг: зафиксируем все  $s_i$  и  $c_q$ , вычислим ожидания скрытых переменных  $z_{iq}$  как

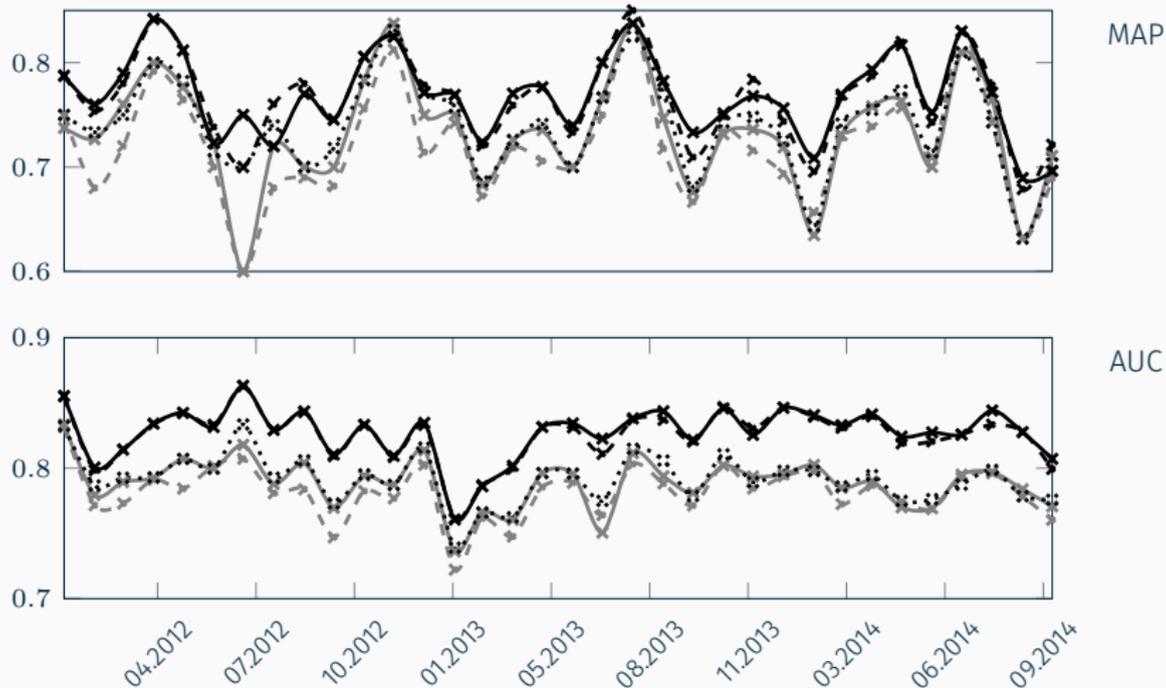
$$\mathbb{E}[z_{iq}] = \begin{cases} 0, & \text{если } x_{tq} = 0, \\ p(z_{iq} = 1 | \exists j \in t z_{jq} = 1) = \frac{\sigma(\mu + s_i + c_q)}{1 - \prod_{j \in t} (1 - \sigma(\mu + s_j + c_q))}, & \text{если } x_{tq} = 1; \end{cases}$$

- M-шаг: зафиксируем  $\mathbb{E}[z_{iq}]$ , обучим логистическую модель

$$\mathbb{E}[z_{iq}] \sim \sigma(\mu + s_i + c_q).$$

# РЕЗУЛЬТАТЫ

- Ну и, конечно, работает хорошо.



РЕЙТИНГ-ЛИСТ ИГРОКОВ НА ЯНВАРЬ 2015

Показывать по  записей

Введите id или начало фамилии:

Место	id	Фамилия	Имя	Отчество	Команда	Сыграно	Взято	Рейтинг
1	27177	Ромашова	Вероника	Михайловна	ЛКИ	5278	3784	545.753
2	3083	Белявский	Дмитрий	Михайлович	ЛКИ	4831	3396	543.520
3	27403	Руссо	Максим	Михайлович	ЛКИ	7180	5205	534.540
4	4270	Брутер	Александра	Владимировна	ЛКИ	8244	5974	533.405
5	18332	Либер	Александр	Витальевич	Рабочее название	8658	6210	532.921
6	1585	Архангельская	Юлия	Сергеевна	Ксеп	7542	5286	531.866
7	24384	Пашковский	Евгений	Александрович	ЛКИ	7552	5374	531.327
8	8333	Губанов	Антон	Александрович	Команда Губанова	4475	3153	530.871
9	16332	Крапиль	Николай	Валерьевич	Ксеп	6927	4899	530.559
10	21487	Моносов	Борис	Яковлевич	Команда Губанова	5115	3645	530.175

Рейтинг ЧГК

ИГРОКИ

КОМАНДЫ

ТУРНИРЫ

ВОПРОСЫ

Александр Друзь vs. Максим Поташев

Ссылка на сайт рейтинга МАК

**Друзь**  
Александр Абрамович



**Поташев**  
Максим Оскарович

Ссылка на сайт рейтинга МАК

ИСТОРИЯ РЕЙТИНГА

ИСТОРИЯ ТУРНИРОВ

ИСТОРИЯ РЕЙТИНГОВ

Рейтинг	Команда	Место	Дата	Место	Команда	Рейтинг
450.138	Трансфера	303	Январь 2015	31	Афина	514.643
446.669	Трансфера	336	Декабрь 2014	29	Афина	514.697
442.559	Трансфера	348	Ноябрь 2014	25	Афина	513.155
437.999	Трансфера	400	Октябрь 2014	25	Афина	512.083
445.791	Трансфера	306	Сентябрь 2014	23	Афина	515.737
441.473	Трансфера	372	Август 2014	23	Афина	516.513
446.730	Трансфера	284	Июль 2014	22	Афина	516.755
448.598	Трансфера	322	Июнь 2014	24	Афина	516.603

Спасибо за внимание!