



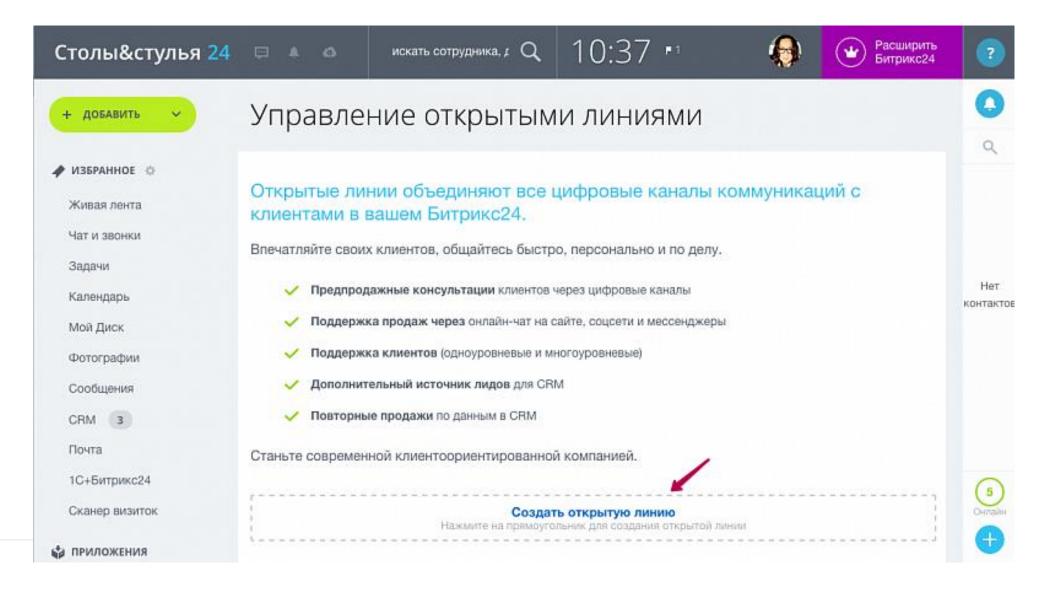


- У вас есть какой-нибудь интернет-магазин
- Нужно скоммуницировать экспертов интернет-магазина с покупателями из разных каналов...
- Люди начинают задавать вопросы...
- Ваши сотрудники начинают им отвечать...

#### Как работают Открытые линии

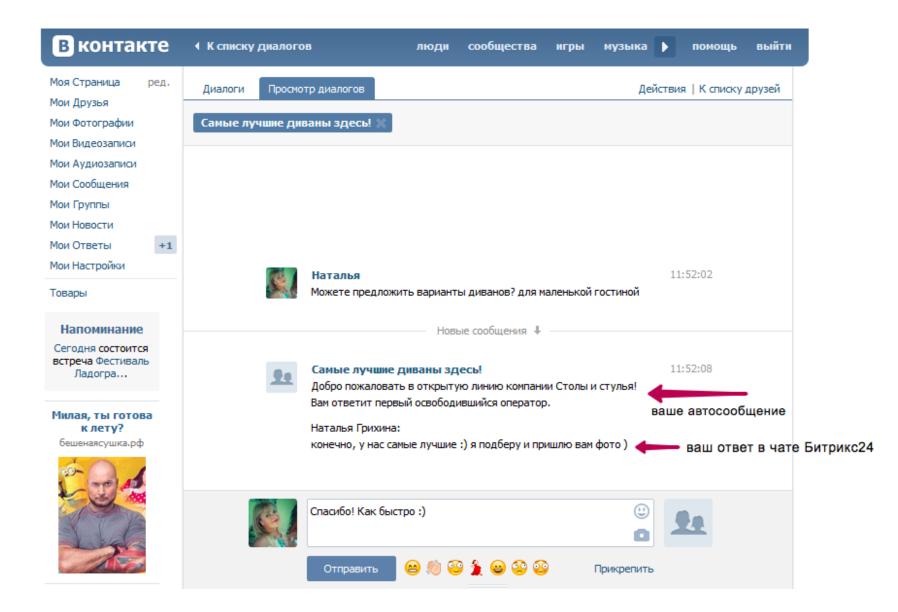




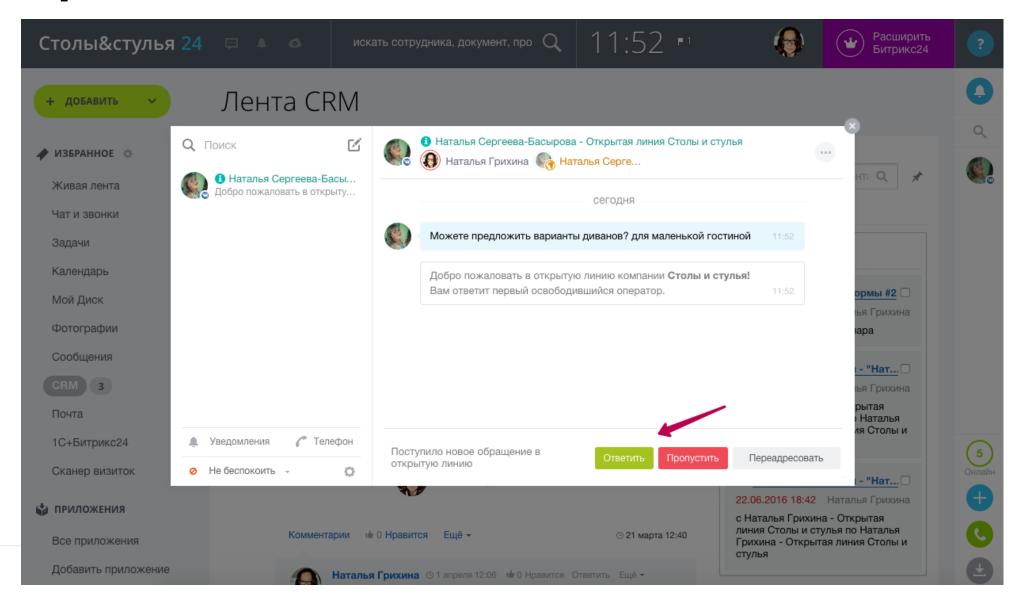




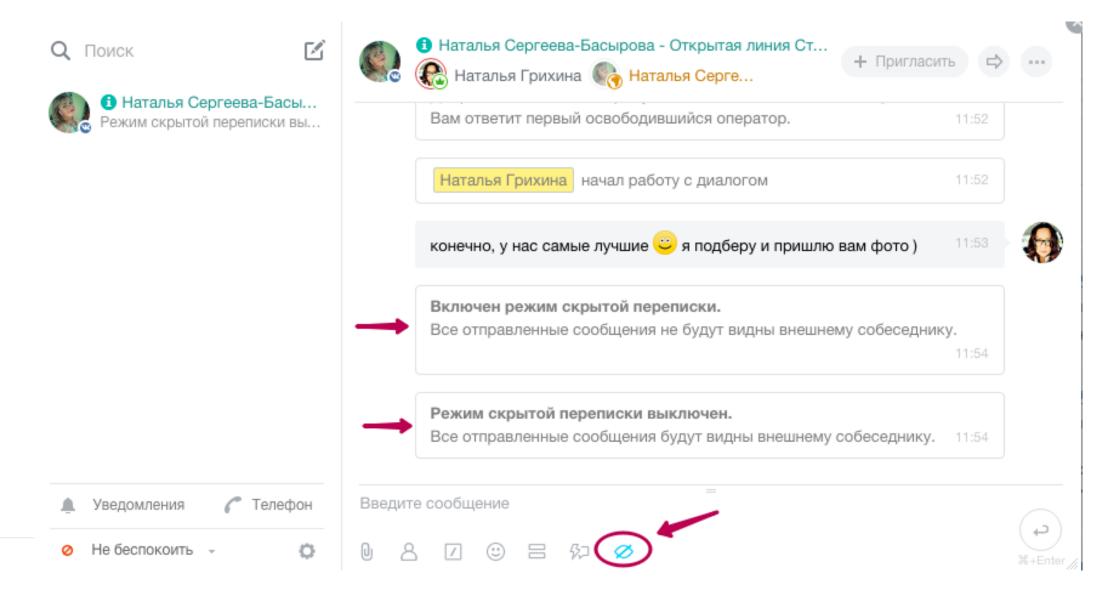
◀ Назад













### Призовем Ктулху?

- Клиент состыкован с менеджером
- Но клиенты задают часто похожие вопросы
- И получают похожие ответы
- База знаний, ЧАВО...
- А вдруХ АІ поможет?

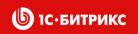




## Все готовы к отжигу в печи?

- Начнет болеть голова
- Из носа потечет кровь
- Пропадет слух
- Возможны судорги и левитация...





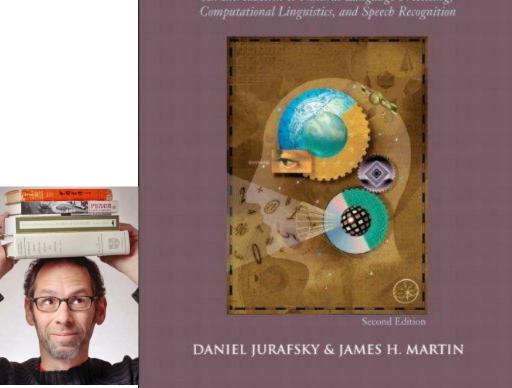
SPEECH AND

LANGUAGE PROCESSING

#### Здравствуй, NLP

- Регулярки? Да! ©
- Куча экспертов и ... регулярки?
- Предикативная логика и базка?
- Семантические графы и ЛСД

 $(C_{20}H_{25}N_3O)$ 





## Национальный корпус русского языка и ... «тайные знания»

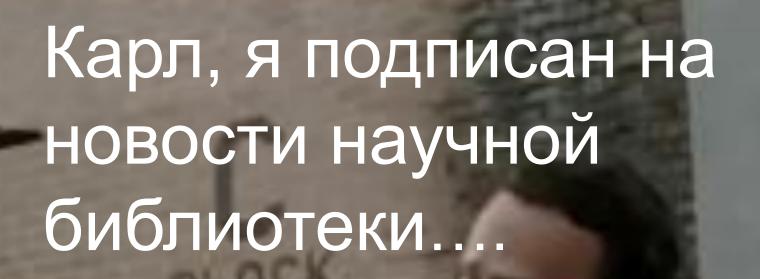
- ruscorpora.ru
- Томита-парсер (Яндекс)
- Морфологический словарь
- Словарь синонимов
- Тезаурус/семантическая сеть
- «Тайные знания» Яндекс,

ABBYY...

Синсеты в WordNet связаны между собой различными семантическими отношениями:

- гипероним (breakfast → meal) (завтрак → прием пищи);
- гипоним (meal  $\rightarrow$  lunch) (прием пищи  $\rightarrow$  обед);
- has-member (faculty → professor) (факультет → профессор);
- member-of (pilot → crew) (пилот → экипаж);
- мероним: has-part (table → leg) (стол → ножка);
- антоним (leader  $\rightarrow$  follower) (лидер  $\rightarrow$  последователь).



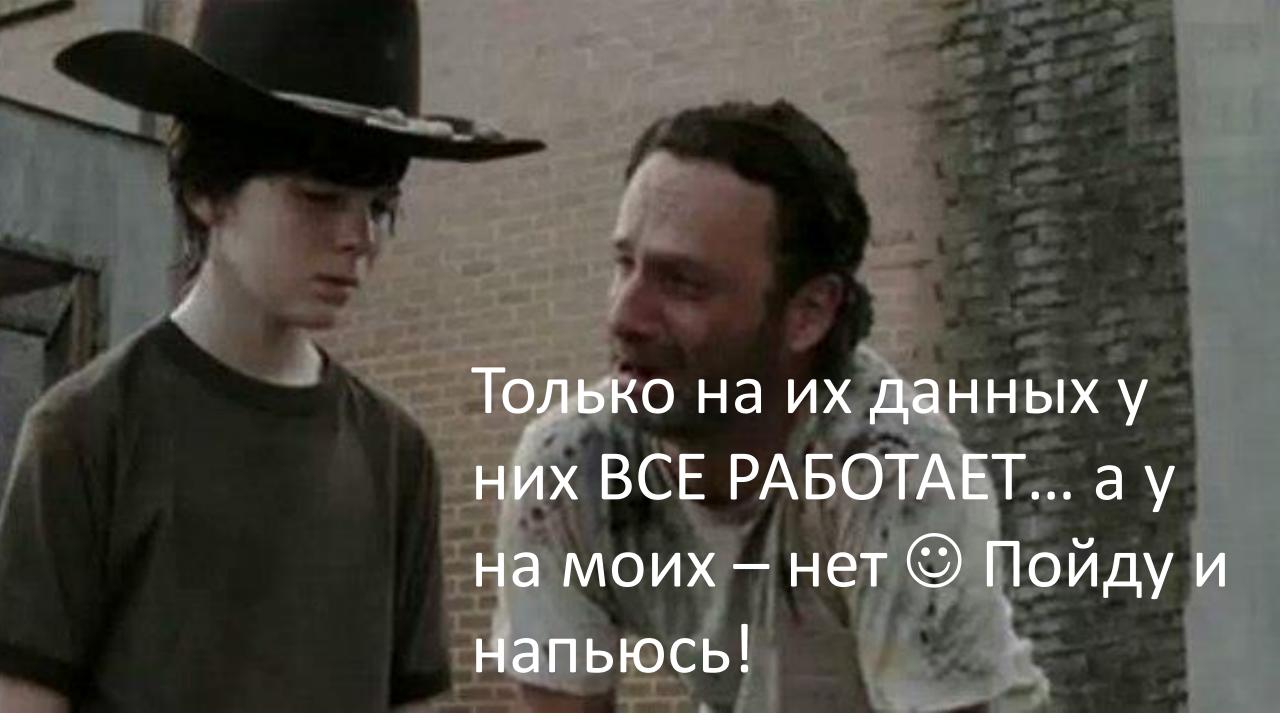


Это очень круто, пап!



Cornell University Library

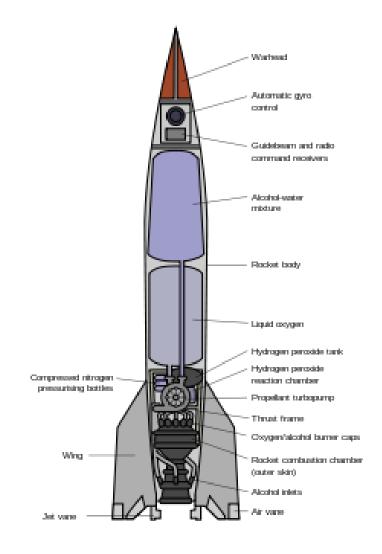
arXiv.org





#### «Из грязи» ... в князи!

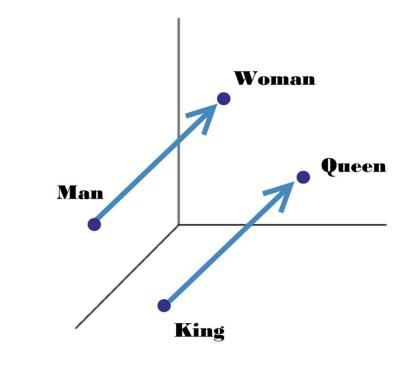
- Зачем изучать линейную алгебру?
- Зачем понимать логистическую регрессию?
- Зачем знать модели <del>Макарова</del> Маркова?
- Айда сразу в ... Deep Learning!

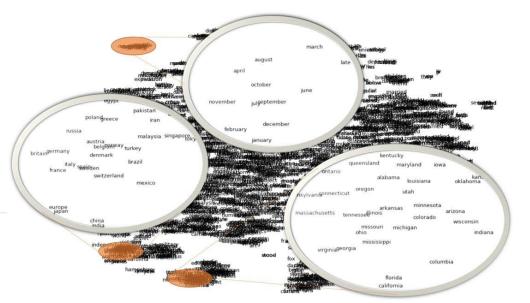




# Нам поможет «черная магия» нейронок

- DSSM (Microsoft) и другие...
- Яндекс «Палех», Google SmartReply
- Word2Vec, GloVe



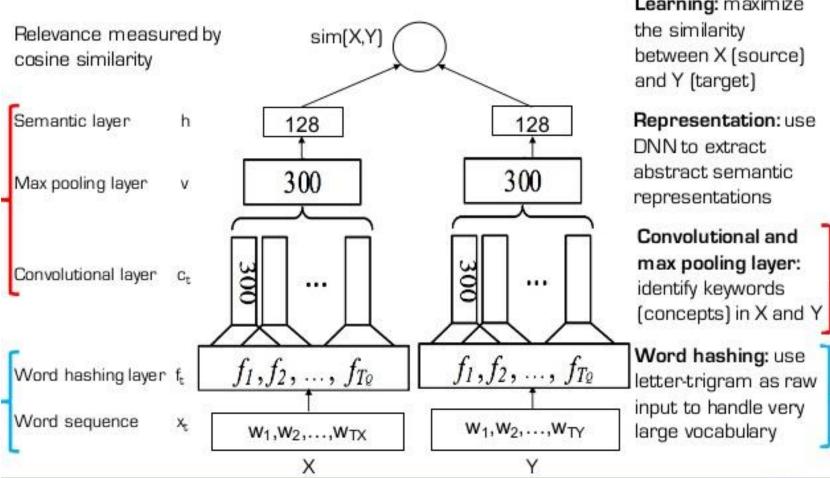




#### Суть архитектуры на пальцах...

- 3 (N)-граммы, tfidf
- Сжатие размерности (нелинейное)
- Метрика похожести какая «НИТЬ», например скалярное произведение векторов: arxiv (1508.01585v2)

#### **DSSM** schema with CNN



Learning: maximize



### 

- Многое зависит от ваших данных
- Их объема, структуры
- Фазы Луны
- Года по китайскому календарю
- Дня и месяца рождения руководителя проекта!

Dev	Test1	Test2	Description
58.2	57.8	53.6	cosine: $k(x,y) = \frac{xy^{T}}{\ x\  \ y\ }$
58.5	57.1	53.3	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d$ , $\gamma = 0.5, d = 2, c = 1$
56.8	54.6	52.6	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d, \ \gamma = 1.0, d = 2, c = 1$
55.0	53.6	48.2	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d, \ \gamma = 1.5, d = 2, c = 1$
57.1	53.7	51.5	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d, \ \gamma = 0.5, d = 3, c = 1$
55.3	52.4	48.7	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d, \ \gamma = 1.0, d = 3, c = 1$
52.5	51.0	47.2	polynomial: $k(x, y) = (\gamma x y^{T} + c)^d, \ \gamma = 1.5, d = 3, c = 1$
61.3	59.9	57.0	sigmoid: $k(x, y) = tanh(\gamma xy^{T} + c), \gamma = 0.5, c = 1$
61.6	60.2	57.1	sigmoid: $k(x, y) = tanh(\gamma xy^{T} + c), \gamma = 1.0, c = 1$
60.2	60.2	55.7	sigmoid: $k(x,y) = tanh(\gamma xy^{T} + c), \gamma = 1.5, c = 1$
60.0	60.3	54.7	RBF: $k(x, y) = exp(-\gamma   x - y  ^2), \gamma = 0.5$
60.2	57.0	54.4	RBF: $k(x, y) = exp(-\gamma   x - y  ^2), \gamma = 1.0$
58.4	57.3	53.8	RBF: $k(x, y) = exp(-\gamma   x - y  ^2), \gamma = 1.5$
60.8	60.3	57.0	euclidean: $k(x,y) = \frac{1}{1+  x-y  }$
42.2	42.5	38.2	exponential: $k(x,y) = exp(-\gamma   x-y  _1), \gamma = 0.5$
41.4	39.5	36.0	exponential: $k(x,y) = exp(-\gamma   x-y  _1), \gamma = 1.0$
48.2	45.1	41.6	exponential: $k(x,y) = exp(-\gamma   x-y  _1), \gamma = 1.5$
51.0	49.5	46.4	manhattan: $k(x,y) = \frac{1}{1 + \ x - y\ _1}$
62.5	61.4	59.0	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{T}+c))}, \gamma = 0.5, c = 1$
62.9	62.1	59.3	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 0.5, c = 1$ GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, c = 1$
62.6	62.1	59.2	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{T}+c))}, \gamma = 1.5, c = 1$
63.1	61.9	58.2	AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{T}+c))}, \ \gamma = 0.5, c = 1$
63.4	61.7	58.7	GESD: $k(x,y) = \frac{1+  x-y  }{1+  x-y  } \cdot \frac{1+exp(-\gamma(xy^{\intercal}+c))}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.5, c = 1$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 0.5, c = 1$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, c = 1$
62.8	62.0	57.7	AESD: $k(x,y) = \frac{1+  x-y  }{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{T}+c))}, \gamma = 1.0, c = 1$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{T}+c))}, \gamma = 1.5, c = 1$
63.5	62.5	60.2	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{T}+c))}, \gamma = 1.0, 2000 \text{ filters}$
64.3	65.1	61.0	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+  x-y  } \cdot \frac{1}{1+  x-y  }$ , $\gamma = 1.0$ , 3000 filters
65.4	65.3	61.0	GESD: $k(x,y) = \frac{1+  x-y  }{1+  x-y  } \cdot \frac{1+exp(-\gamma(xy^{\dagger}+c))}{1+exp(-\gamma(xy^{\dagger}+c))}, \gamma = 1.0, 4000 \text{ filters}$
64.5	62.7	60.1	AESD: $k(x,y) = \frac{1+  x-y  }{0.5} + \frac{1+exp(-\gamma(xy)+c))}{0.5}, \gamma = 1.0.$ 2000 filters
64.3	63.3	62.2	AESD: $k(x,y) = \frac{1+  x-y  }{0.5} + \frac{1+exp(-\gamma(xy)+c))}{0.5}, \gamma = 1.0, 3000 \text{ filters}$
63.9	64.5	61.1	GESD: $k(x,y) = \frac{1}{1+  x-y  } \cdot \frac{1}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, 4000 \text{ filters}$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, 2000 \text{ filters}$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, 3000 \text{ filters}$ AESD: $k(x,y) = \frac{0.5}{1+  x-y  } + \frac{0.5}{1+exp(-\gamma(xy^{\intercal}+c))}, \gamma = 1.0, 4000 \text{ filters}$

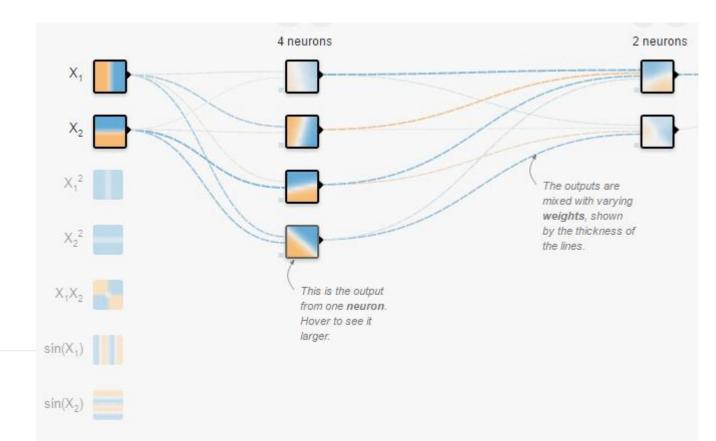


#### На чем «кодить», production

- Тензоры... numpy, nd4j
- Python: Theano, Keras,

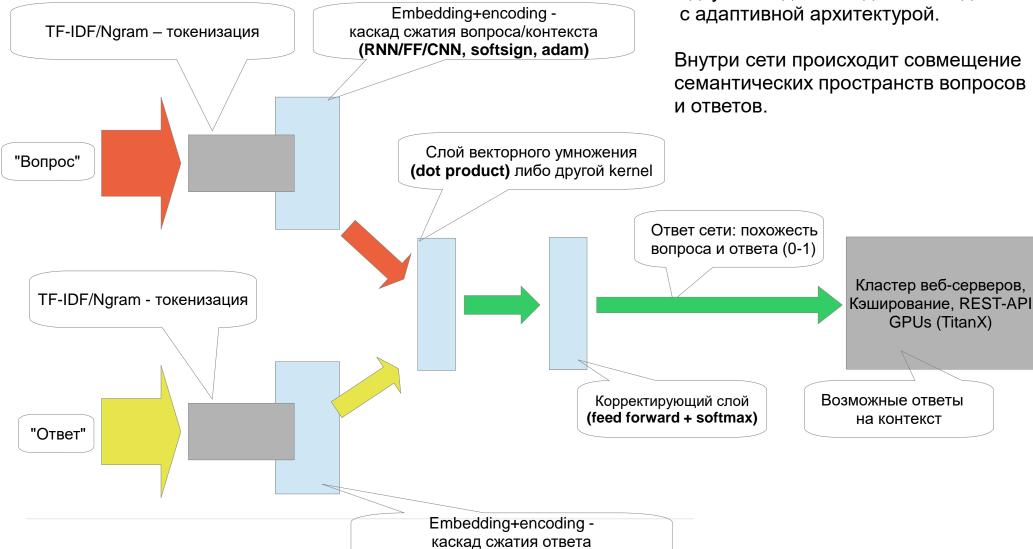
**TensorFlow** 

- Lua: Torch7
- Java/Scala: Deeplearning4j
- Production...





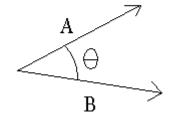
#### «Нейробот» ©



(RNN/FF/CNN, softsign, adam)

Глубокая нейронная сеть с двумя входами и одним выходом

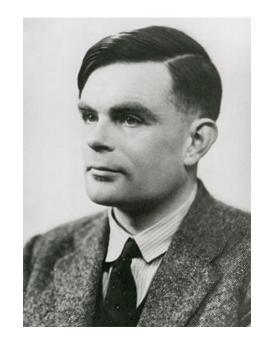
$$A \cdot B = |A| |B| \cos \Theta$$





### Turing Тюнинг...

- Если мало данных... нейронка «бредит»
- Метрики качества: recall@N, уровень «бреда» и т.п.
- Обратный индекс
- Кластеризация похожих ответов
- Многопоточность
- Онлайн-дообучение
- Экстремальный highload
- «Экспериментируем» с мэрией Москвы



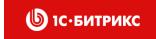




#### Бот-платформа Битрикс24

- https://dev.1c-bitrix.ru/learning/course/?COURSE\_ID=93
- Приглашаем дата-<del>сатанистов</del> сайнтистов к сотрудничеству
- Пишите NLP + Deep + Convolution+ Adversarial + Autoencoder + Neural Turing Machine + ... чат-боты
- Всем отличного настроения и качественных моделей! ©





Спасибо за внимание! Вопросы?

### Александр Сербул

@AlexSerbul

Alexandr Serbul serbul@1c-bitrix.ru

