

# Identifying Comparative Questions on the Russian Web

Alexander Bondarenko

Referees:

Prof. Dr. Benno Stein

Dr. rer. nat. habil. Andreas Jakoby

Advisors:

Prof. Dr. Matthias Hagen

M.Sc. Michael Völske

# Motivation

## Variety of choices



# Motivation

## Variety of choices



# Motivation

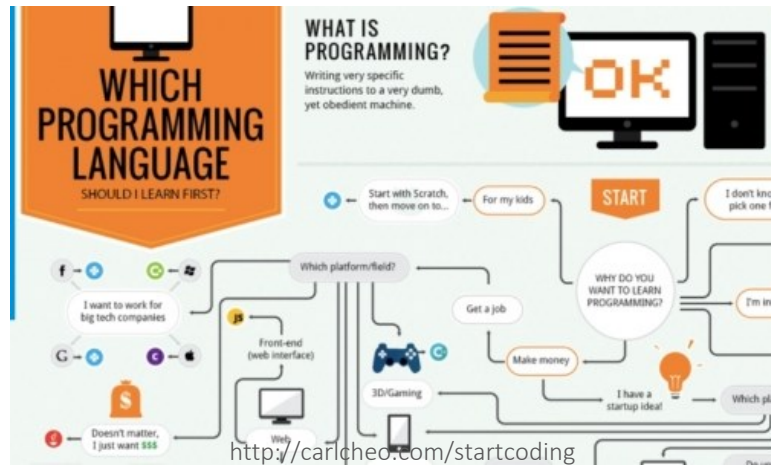
## Variety of choices





# Motivation

## Variety of choices



**So many choices  
So little time!**



# Motivation

Which to buy XX or YY?



# Motivation

Which to buy XX or YY?



Ask your friend

# Motivation

Which to buy XX or YY?



Community  
Q & A



Ask your friend



# Motivation

Which to buy XX or YY?



Community  
Q & A



Ask your friend



"OK Google"

Search engine



# Motivation

The screenshot shows the Google Trends interface for the search query "difference between iphone 8 and x". The "Google Trends" logo is circled in orange. The search results are categorized into "Related topics" and "Related queries".

**Related topics**

Rank	Topic	Status
1	iPhone - Smartphone	Breakout
2	Apple - Technology company	Breakout
3	iPhone 6 - Topic	Breakout
4	Camera - Topic	Breakout
5	Verizon Wireless - Telecommunications comp...	Breakout

**Related queries**

Rank	Query	Status
1	the difference between iphone 8 and iphone x	Breakout
2	difference between the iphone 8 and x	Breakout
3	what is difference between iphone 8 and ip...	Breakout
4	what is the difference between iphone 8 and x	Breakout
5	what is the difference between the iphone 8 an...	Breakout

An orange arrow points from the word "Questions" to the third query in the "Related queries" list, which is also circled in orange.

# Comparative queries are not treated differently from the others

Google

difference between iphone 7 and 8

All Images Videos Shopping News More Settings Tools


About 123.000.000 results (0,55 seconds)

**iPhone 8 Vs iPhone 7: What's The Difference? - Forbes**  
<https://www.forbes.com/sites/.../apple-iphone-8-vs-iphone-7-whats-the-difference/> ▼  
Sep 18, 2017 - did not follow the iPhone 7 with an 'iPhone 7S' so instead we have the iPhone 8 and iPhone 8 Plus as well as a radically redesigned (and jaw droppingly expensive) all new iPhone X. So what is the difference between the iPhone 8 and iPhone 7 and is it a worthy upgrade? Here's what you need to know...  
You visited this page on 1/29/18.

**iPhone 8 Vs iPhone 7 Vs iPhone 6S Vs iPhone 6: What's The ... - Forbes**  
<https://www.forbes.com/.../apple-iphone-8-vs-iphone-7-vs-iphone-6s-vs-iphone-6-wh...> ▼  
Jan 12, 2018 - This guide will explain the important differences between the iPhone 6 (2014), iPhone 6S (2015), iPhone 7 (2016) and iPhone 8 (2017) - and discuss whether you should upgrade. Here's everything you need to know... Apple. Left to right: iPhone 8, iPhone 7, iPhone 6S, iPhone 6. Design - Fluctuations In ...

**iPhone 8 vs iPhone 7: What's changed? | Trusted Reviews**  
[www.trustedreviews.com](http://www.trustedreviews.com) > News ▼  
iPhone 7 vs iPhone 8 Design: What's the difference? Apple played it safe with the iPhone 7 last year, deciding to keep the overall form-factor and size the same as both the iPhone 6S and 6. And again this time around there isn't a whole lot different – all the big changes seem to have gone into the iPhone X. Instead, the ...

Shop for difference between ip... Sponsored ⓘ

 Apple iPhone 7 32GB Schwarz [11,94cm (4,7") Retina HD Display, iOS 10, A10,...  
**€589,00**  
notebooksbilliger.de  
Free shipping  
By Google

# Comparative queries are not treated differently from the others

Яндекс

разница между айфон 7 и 8



Найти

Создать аккаунт

ПОИСК КАРТИНКИ ВИДЕО КАРТЫ МАРКЕТ НОВОСТИ ПЕРЕВОДЧИК ЕЩЁ

## Чем отличается iPhone 8 от iPhone 7

[iphones.ru](#) > [iNotes/740826](#) ▼

Но свою семерку я вряд ли поменяю, смысла не вижу, **разница** незначительная. Да и знаю этот прикоп с новым процессором, с 7 так же было – пока разработчики подтянутся, выйдет новый **айфон** с новым улучшенным процессором.

Нашёлся 101 млн результатов

1 013 показов в месяц

[Дать объявление](#)

## 24 Сравнение iPhone 7 и iPhone 8: в чем разница и стоит ли...

[24hitech.ru](#) > [sравnenie-iphone-7-i-iphone-8...raznica...](#) ▼

...По причине того, что все основные эволюционные шаги последних моделей **iPhone** касались исключительно 5,5-дюймовых фаблетов с приставкой "Plus", отыскать существенные **отличия** между **iPhone 7** и 8 довольно сложно, однако **разница** в...

## Айфон 8 и Айфон 7 сравнение, отличие двух устройств

[kazizilkree.ru](#) > [ajfon-8-i-ajfon-7-sravnenie/](#) ▼

Вот мы решили сравнить **Айфон 8/8+** и **айфон 7/7+** с ключевыми особенностями, чтобы дать пользователям представление о **разнице** между **iPhone 8/8+** и **iPhone 7/7+**, прежде чем купить их. **Айфон 8** и **айфон 7** сравнение двух девайсов...

# Possible search result page representation for the comparative query

All

Images

Videos

Shopping

News

More

Settings

Tools

About 123.000.000 results (0,55 seconds)

iPhone 8 Vs iPhone 7: What's The Difference? - Forbes

<https://www.forbes.com/sites/.../apple-iphone-8-vs-iphone-7-whats-the-difference/>

Sep 18, 2017 - did not follow the iPhone 7 with an 'iPhone 7S' so instead we have the iPhone 8 and iPhone 8 Plus as well as a radically redesigned (and jaw droppingly expensive) all new iPhone X. So what is the difference between the iPhone 8 and iPhone 7 and is it a worthy upgrade? Here's what you need to know...

You visited this page on 1/29/18.

	iPhone 8	iPhone 7
iOS	iOS 11	iOS 11 update available
Colours	Silver, Gold, Space Grey	Silver, Gold, Rose Gold, Black, Jet Black
Display	4,7in, 1334x750 IPS, True Tone, 326ppi	4,7in, 1334x750 IPS, 326ppi
Processor	Apple A11 Bionic, M11 co-processor	A10 Fusion, M10 co-processor
Storage	64GB/256GB	32GB/128GB
Rear camera	12MP, f/1.8 with OIS	12Mp, f/1.8 with OIS
Front camera	7Mp, f/2.2	7Mp, f/2.2

iPhone 8 vs iPhone 7: What's changed? | Trusted Reviews

[www.trustedreviews.com](http://www.trustedreviews.com) > News

iPhone 7 vs iPhone 8 Design: What's the difference? Apple played it safe with the iPhone 7 last year, deciding to keep the overall form-factor and size the same as both the iPhone 6S and 6. And again this time around there isn't a whole lot different – all the big changes seem to have gone into the iPhone X. Instead, the ...

Shop for difference between ip...

Sponsored

Apple iPhone 7 32GB Schwarz [11,94cm (4,7") Retina HD Display, iOS 10, A10,...

€589,00

[notebooksbilliger.de](http://notebooksbilliger.de)

Free shipping

By Google

Tabular  
summary

	iPhone 8	iPhone 7
iOS	iOS 11	iOS 11 update available
Colours	Silver, Gold, Space Grey	Silver, Gold, Rose Gold, Black, Jet Black
Display	4,7in, 1334x750 IPS, True Tone, 326ppi	4,7in, 1334x750 IPS, 326ppi
Processor	Apple A11 Bionic, M11 co-processor	A10 Fusion, M10 co-processor
Storage	64GB/256GB	32GB/128GB
Rear camera	12MP, f/1.8 with OIS	12Mp, f/1.8 with OIS
Front camera	7Mp, f/2.2	7Mp, f/2.2

Bauhaus-Universität Weimar

Alexander Bondarenko - Identifying comparative questions on the Russian web

13



We need to be able to recognize  
a user's intent for comparison



# Study research questions

RQ1: What are the strong textual signals that distinguish comparative questions from other questions?

RQ2: Can we build an effective classifier to automatically recognize comparative questions on the Russian web?

RQ3: How often and in what categories do users ask comparative questions on the Russian web?

# Different types of user queries to the search engine

## Query

difference between iphone 7 and iphone 8



# Different types of user queries to the search engine

## Query

difference between iphone 7 and iphone 8



## Question(-like) query

what is the difference between an iphone 7 and iphone 8



# Different types of user queries to the search engine

## Query

difference between iphone 7 and iphone 8



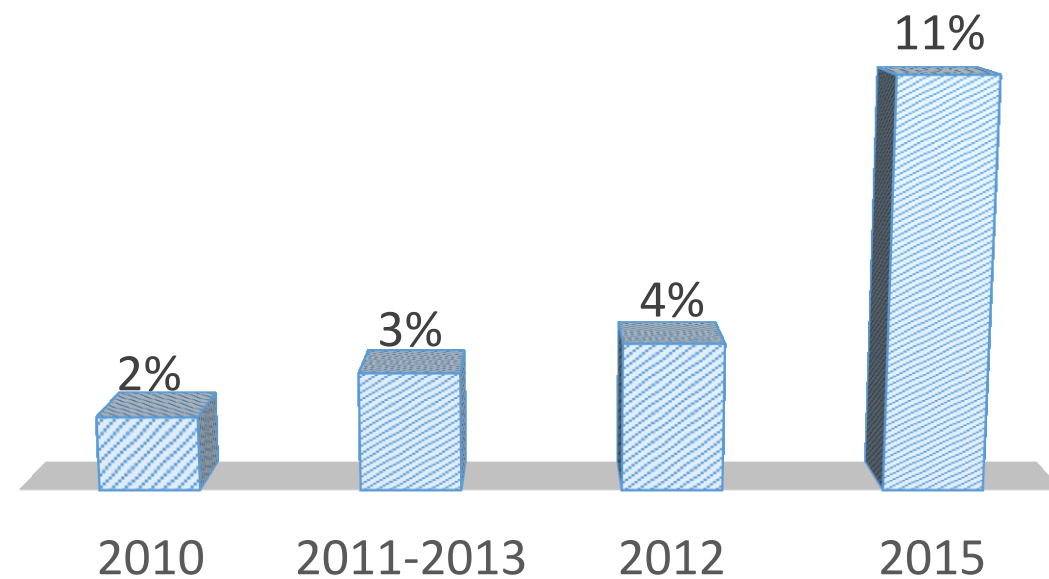
Question(-like) query

what is the difference between an iphone 7 and iphone 8



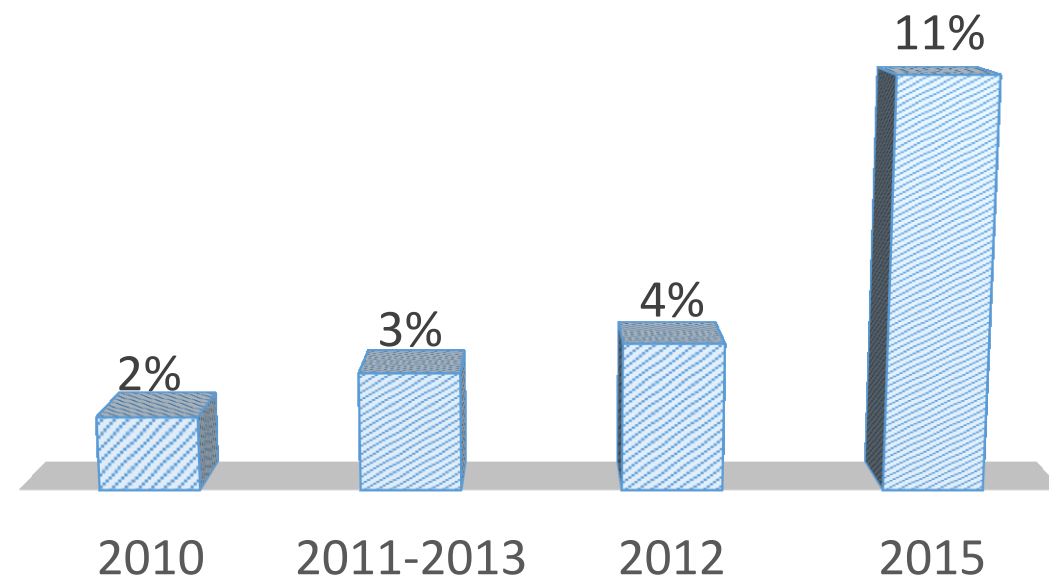


Question queries ~ **2%** of the Web queries  
Yahoo! Research  
[Pang and Kumar, 2011]



Question queries ~ **2%** of the Web queries  
Yahoo! Research  
[Pang and Kumar, 2011]

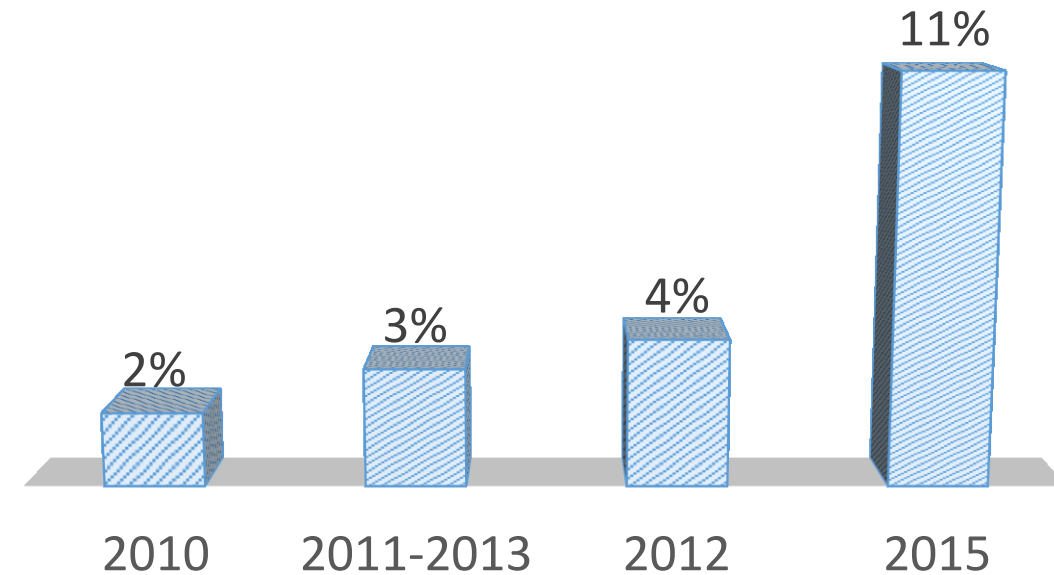
Queries written in natural language ~ **3.2%**  
Microsoft Research  
[ White et al., 2015]



Question queries ~ **2%** of the Web queries  
Yahoo! Research  
[Pang and Kumar, 2011]

Queries written in natural language ~ **3.2%**  
Microsoft Research  
[ White et al., 2015]

Question queries ~ **4%** of the Web queries  
Bauhaus Universität, Ural Federal University  
[Völske et al., 2015]

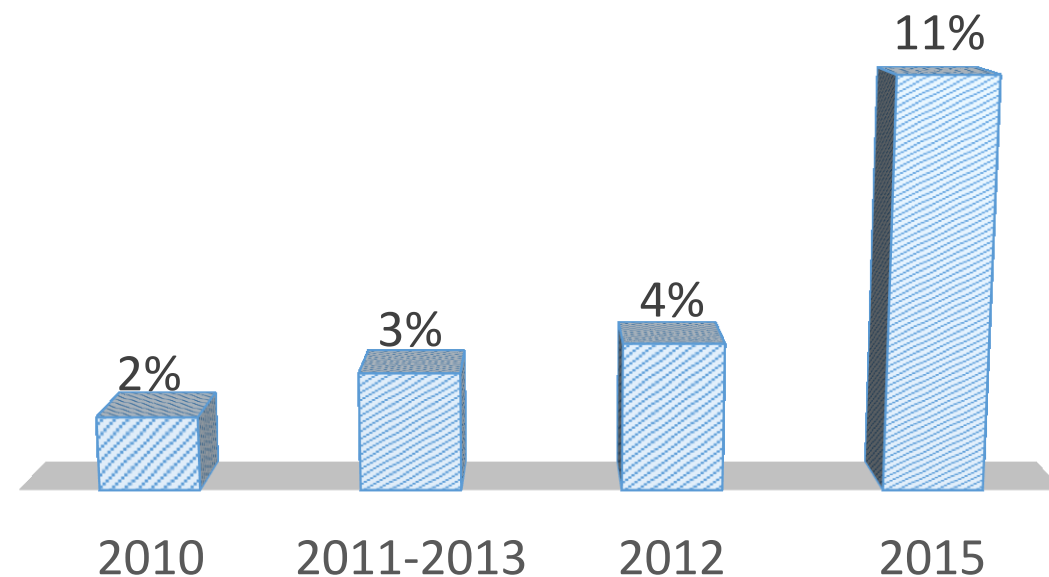


Question queries ~ **2%** of the Web queries  
Yahoo! Research  
[Pang and Kumar, 2011]

Queries written in natural language ~ **3.2%**  
Microsoft Research  
[ White et al., 2015]

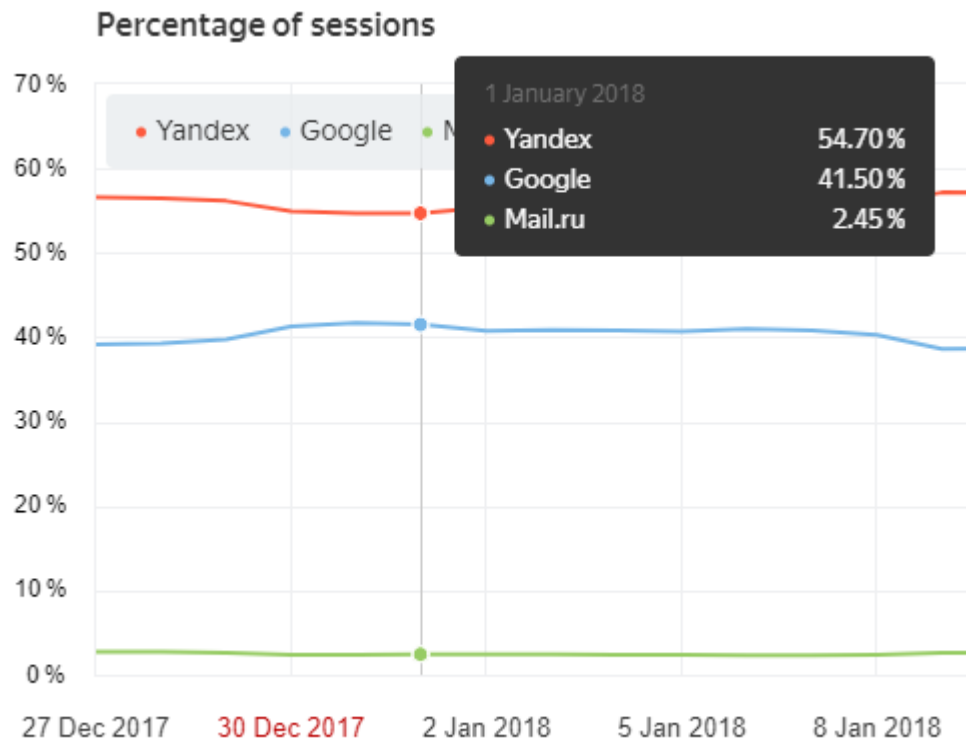
Question queries ~ **4%** of the Web queries  
Bauhaus Universität, Ural Federal University  
[Völske et al., 2015]

Question queries ~ **11%** in the voice search  
Yahoo! Research  
[Guy, 2016]



# Available data

## Yandex



<https://radar.metrka.yandex/>

## ОТВЕТЫ@mail.ru

ОТВЕТЫ@mail.ru Категории + Спросить | Лидеры Поиск по вопросам

Опубликовать вопрос

### Вопросы участников

Открытые На голосовании Лучшие

46 новых вопросов

- как это убрать пожалуйста  
Денис Мухин в "Прочее компьютерное", 1 минуту назад • ответить
- Приветствую. У меня вот такой вопрос есть (в описании)  
sssfse ssefs в "Болезни, Лекарства", 1 минуту назад • ответить
- ЧистАпсихологически.. Есть тут ответчики-с Кому я неприятна.. может бешу.. али тошнит где? Наваяло. Надуло таксказать.  
Фира Фу в "Психология", 1 минуту назад • ответить



# Yandex query logs pre-processing following [Völske et al., 2015]



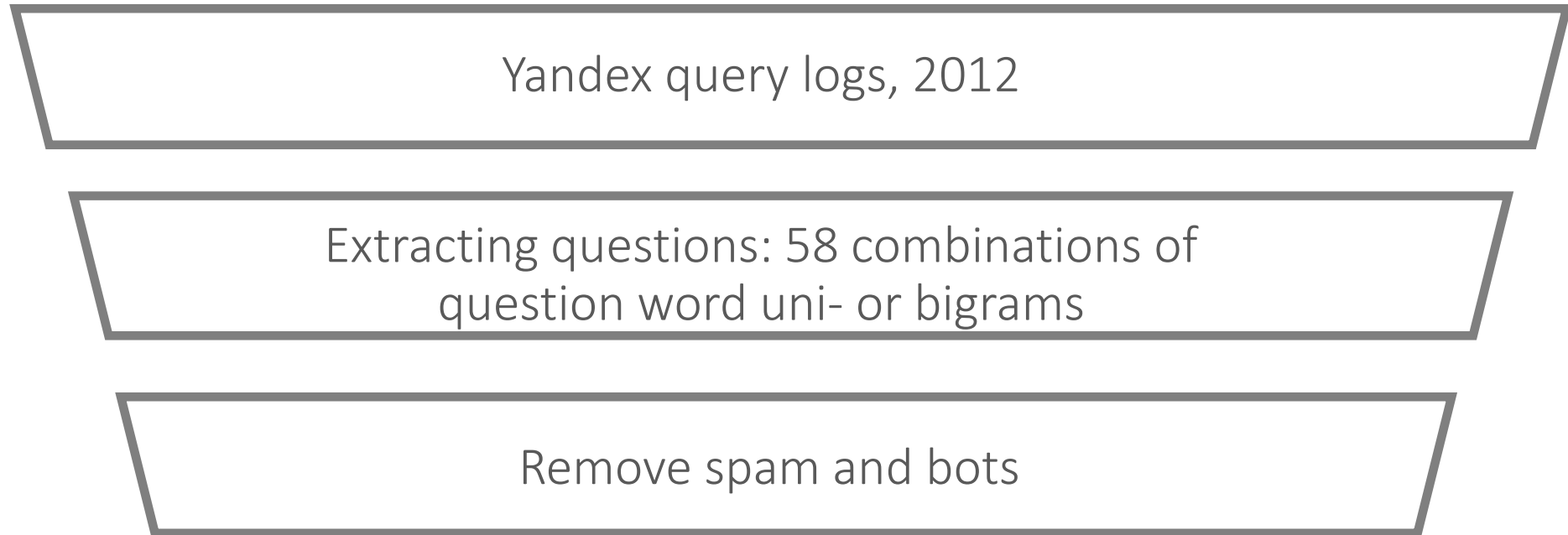
Yandex query logs, 2012

# Yandex query logs pre-processing following [Völske et al., 2015]

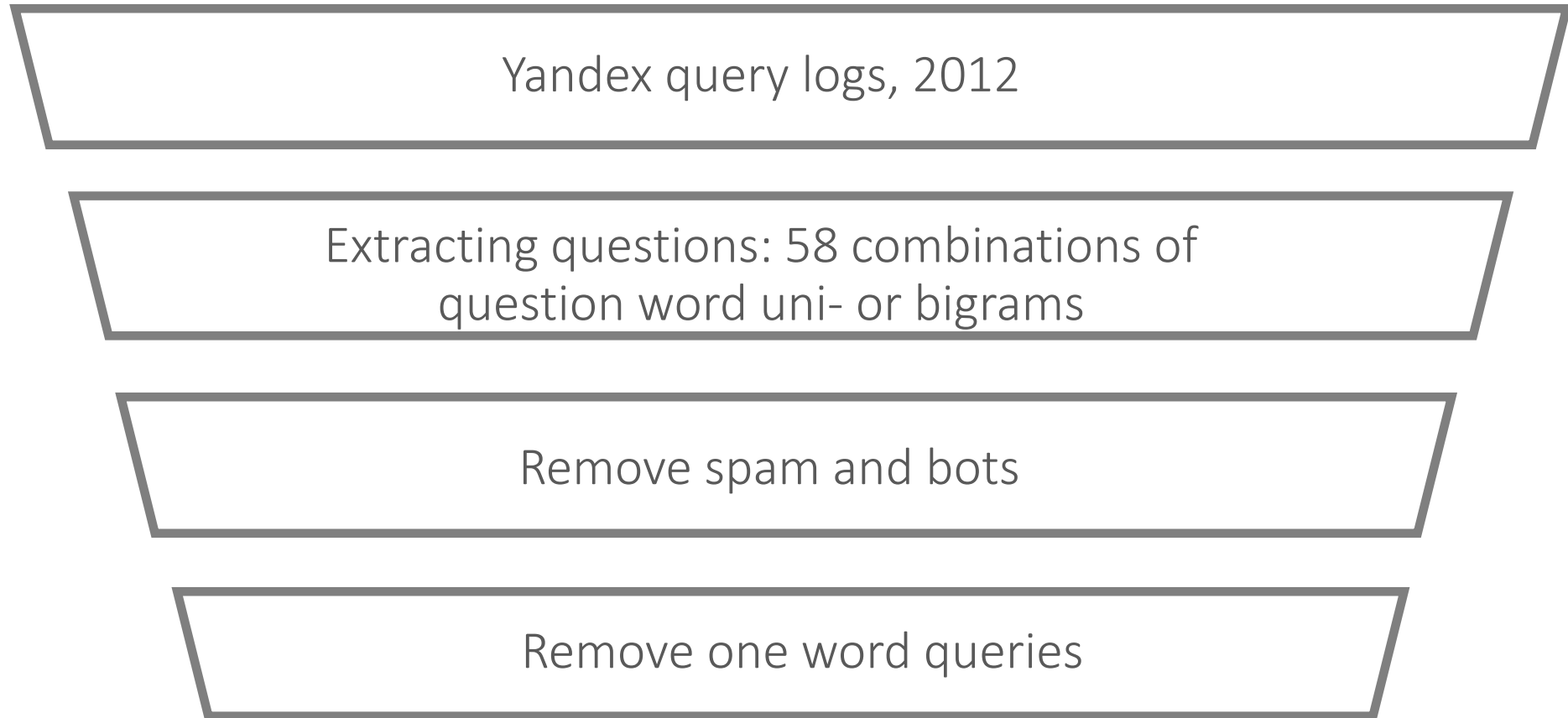
Yandex query logs, 2012

Extracting questions: 58 combinations of  
question word uni- or bigrams

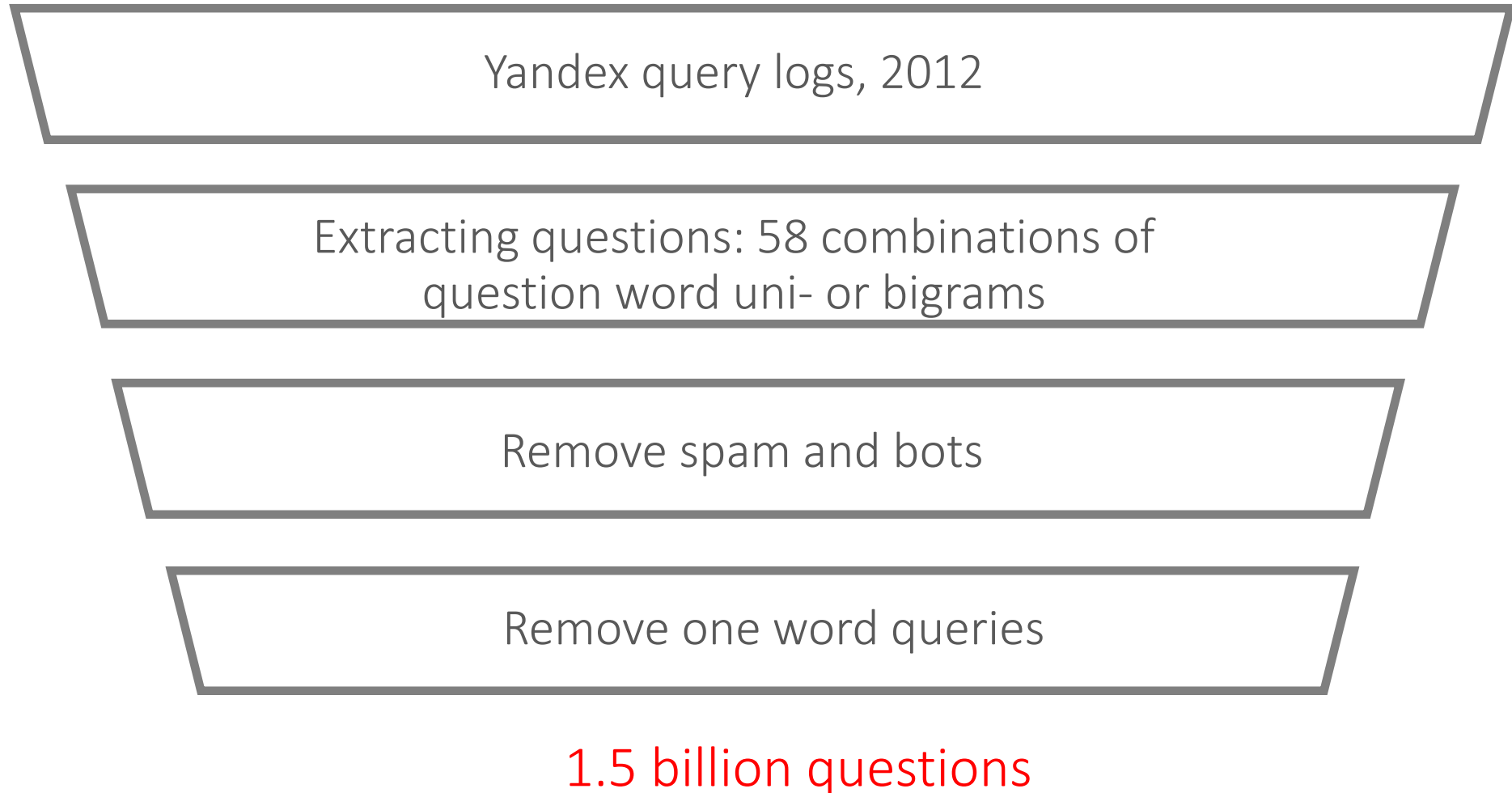
# Yandex query logs pre-processing following [Völske et al., 2015]



# Yandex query logs pre-processing following [Völske et al., 2015]



# Yandex query logs pre-processing following [Völske et al., 2015]





# Example: question – answer on Otvety@Mail.ru

The screenshot shows the Otvety@Mail.ru website interface. The top navigation bar includes links for Mail.Ru, Почта, Мой Мир, Одноклассники, Игры, Знакомства, Новости, Поиск, and Все проекты. The main header features the site logo, navigation links for Категории, Спросить, and Лидеры, and a search bar labeled 'Поиск по вопросам'.

On the left sidebar, there are links for 'Шаблоны для писем в почте', 'Регистрация', and various relationship categories: 'Знакомства, Любовь, Отношения', 'Дружба', 'Знакомства', 'Любовь', 'Отношения', 'Расставания', and 'Прочие взаимоотношения'. At the bottom of the sidebar, there is a section for 'Вопросы-лидеры' with a featured question: 'Как выявить симпатию от парня?' with 1 vote.

The main content area displays a question from user 'Фая Марфуша' (Ученик (26)): 'ктонибудь знает? а в продаже имеется айфон 5? сколько стоит? чем отличается от 4g и какой лучше 4 или 5)'. The question is marked as 'Вопрос решён 6 лет назад'. Below the question are interaction buttons: 'Нравится', 'Ответить', and social media sharing icons.

The 'ЛУЧШИЙ ОТВЕТ' (Best Answer) section shows a response from user 'НиколЯ' (6 лет назад, Оракул (68430)): '5 лучше чем 4...на 1см больше...))'. This answer also has 'Нравится' and 'Пожаловаться' buttons.

At the bottom of the page, a feedback prompt asks: 'Вы нашли ответ на свой вопрос?' with buttons for 'Да, спасибо!' and 'Нет :( )'.

# Otvety@Mail.ru data pre-processing

Otvety@Mail.ru: 11.2 million questions with answers

```
'best': {'kpd': 14.39, 'email': '1nlo.1@mail.ru', 'nick': 'НиколЯ', 'qid': '70856473', 'comcnt': '0', 'vip': 0, 'rating': '2', 'id': '379029027', 'points': '54333', 'source': '', 'usrid': '9218636', 'lvl': 'Оракул', 'added': '2012-02-06 22:16:32', 'atext': '5 лучше чем 4...на 1см больше...))'}, 'qtext': 'ктонибудь знает? а в продаже имеется айфон 5 ? сколько стоит? чем отличается от 4g и какой лучше 4 или 5)', 'comcnt': '0', 'cid': '1335', 'vip': 0, 'cancomment': '0', 'id': '70856473', 'points': '28', 'rating': '0', 'arating': '3', 'lvl': 'Ученик', 'added': '2012-02-06 22:14:43', 'kpd': 15.38, 'email': 'kashapchik93@mail.ru', 'nick': 'Фая Марфуша', 'anscnt': '3', 'polltype': '', 'qcomment': '', 'usrid': '35975834', 'category': {'name': 'Любовь', 'urlname': 'loving'}, 'answers': [{'kpd': 15.38, 'email': 'darkness123123@mail.ru', 'nick': 'Michael Myerss', 'qid': '70856473', 'comcnt': '0', 'vip': 0, 'rating': '1', 'id': '379029523', 'points': '40', 'source': 'http://iphone5news.ru/?p=5587', 'usrid': '62123520', 'lvl': 'Ученик', 'added': '2012-02-06 22:18:37', 'atext': 'айфон 5 еще нет в продаже он появится или в августе или осенью как говорят на английских сайтах айфона !\n'}, {'kpd': 0, 'email': 'kost2001kv@mail.ru', 'nick': 'Кукла', 'qid': '70856473', 'comcnt': '0', 'vip': 0, 'rating': '0', 'id': '379811027', 'points': '204', 'source': '', 'usrid': '47740509', 'lvl': 'Ученик', 'added': '2012-02-10 13:33:05', 'atext': 'там выдвезная клавиатура'}], 'state': 'R'
```

# Otvety@Mail.ru data pre-processing

Otvety@Mail.ru: 11.2 million questions with answers

**'best': {'atext': '5 лучше чем 4...на 1см больше...))'},**  
'379029027', 'points': '54333', 'source': '', 'usrid': '9218636', 'lvl': 'Оракул', 'added': '2012-02-06 22:16:32', 'atext': '5 лучше чем 4...на 1см больше...))', 'qtext': 'кто нибудь знает? а в продаже имеется айфон 5 ? сколько стоит? чем отличается от 4g и какой лучше 4 или 5)', 'anscnt': '0', 'cid': '1335', 'vip': '0', 'comment': '0', 'id': '70856473', 'points': '28', 'rating': '0', 'arating': '3', 'lvl': 'Ученик', 'added': '2012-02-06 22:14:43', 'kpd': '15.38', 'email': 'kashapchik93@mail.ru', 'nick': 'Фая Марфуша', 'anscnt': '3', 'polltype': '', 'qcomment': '', 'usrid': '35975834', 'category': {'name': 'Любовь', 'urlname': 'loving'}, 'answers': [{'kpd': '15.38', 'email': 'darkness123123@mail.ru', 'nick': 'Michael Myers', 'qid': '70856473', 'comcnt': '0', 'vip': '0', 'rating': '1', 'id': '379029523', 'points': '40', 'source': 'http://iphone5news.ru/?p=5587', 'usrid': '62123520', 'lvl': 'Ученик', 'added': '2012-02-06 22:18:37', 'atext': 'айфон 5 еще нет в продаже он появится или в августе или осенью как говорят на английских сайтах айфона !\n'}, {'kpd': '0', 'email': 'kost2001kv@mail.ru', 'nick': 'Кукла', 'qid': '70856473', 'comcnt': '0', 'vip': '0', 'rating': '0', 'id': '379811027', 'points': '204', 'source': '', 'usrid': '47740509', 'lvl': 'Ученик', 'added': '2012-02-10 13:33:05', 'atext': 'там выдвижная клавиатура'}], 'state': 'R'

7.7 million questions

First look into the data



# Different types of comparative questions

## Indirect comparative

Where is it better to go on vacation with children 8 and 10 years old?

## Direct comparative

What is better BMW or Mercedes?

# Different types of comparative questions

## Indirect comparative

Where is it better to go on vacation with children 8 and 10 years old?

## Direct comparative

What is better BMW or Mercedes?

## Reasoning comparative

What is better Sony or Canon, and why?

## Non-reasoning comparative

What is bigger Moscow or Paris?

# Different types of comparative questions

## Indirect comparative

Where is it better to go on vacation with children 8 and 10 years old?

## Direct comparative

What is better BMW or Mercedes?

## Reasoning comparative

What is better Sony or Canon, and why?

## Non-reasoning comparative

What is bigger Moscow or Paris?

Difficult case: What is more reliable BMW or Mercedes?



# Comparative textual patterns collection

Formulated **20** patterns

Based on:

Our own expertise

Linguistics research

Computational linguistics/NLP: sentiment analysis  
and opinion mining

Psychological research



# Comparative textual patterns collection

Formulated **20** patterns

Pattern	Number of hits	Comparative %
comp(ADV and ADJ) + or	5.5 million	97

Regex

Human judges



# Comparative textual patterns collection

Formulated **20** patterns

Pattern	Number of hits	Comparative %
comp(ADV and ADJ) + or	5.5 million	97

'что лучше бмв х5 или инфинити фх35'

'what better BMW X5 or Infinity FX35'



# Comparative textual patterns collection

Formulated **20** patterns

Pattern	Number of hits	Comparative %
comp(ADV and ADJ) + or	5.5 million	97

'что лучше бмв х5 или инфинити фх35'

'what better BMW X5 or Infinity FX35'

'как правильнее дверьми или дверями'

'how more correct doors or dors'



# Comparative textual patterns collection

Formulated **20** patterns

Pattern	Number of hits	Comparative %
comp(ADV and ADJ) + or	5.5 million	97

'что лучше бмв х5 или инфинити фх35'

'what better BMW X5 or Infinity FX35'

'как правильнее дверьми или дверями'

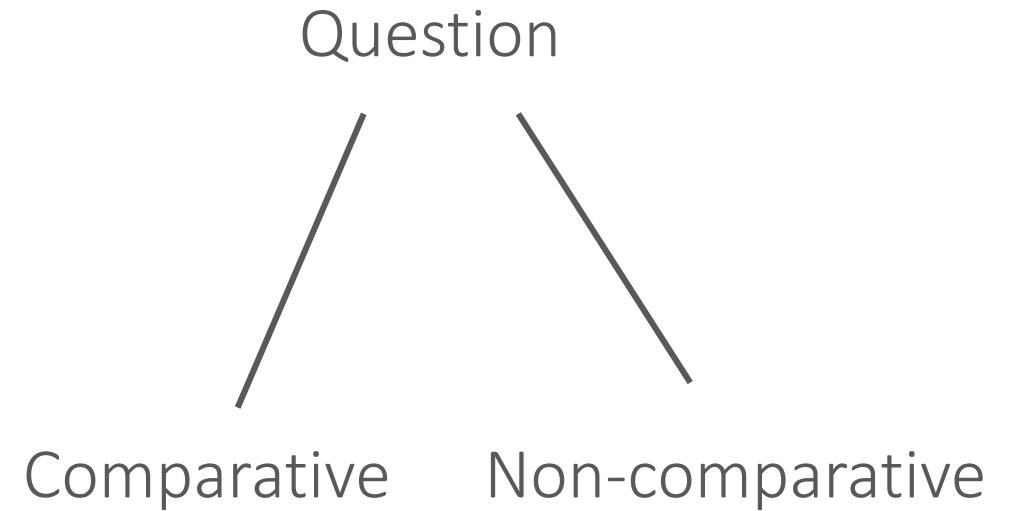
'how more correct doors or dors'



We are interested in queries seeking for reasoning support



We use the collection of the patterns  
to classify questions in our corpora

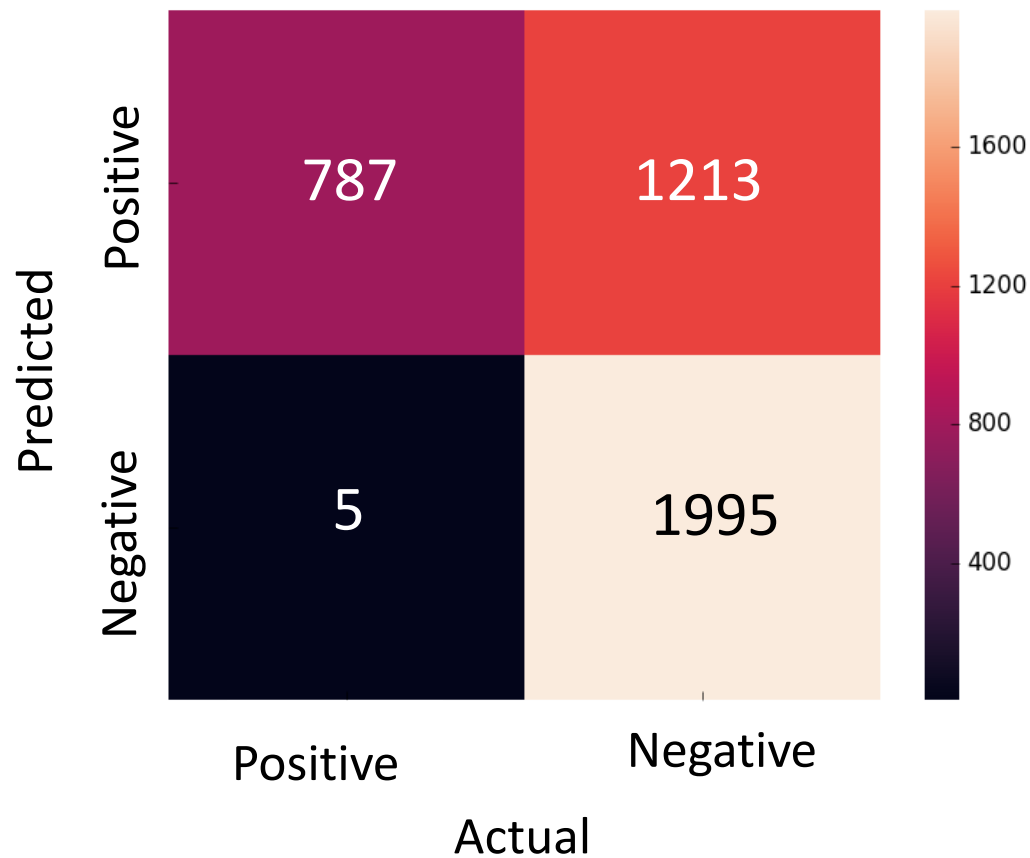


# Pattern-based classification of reasoning comparative questions

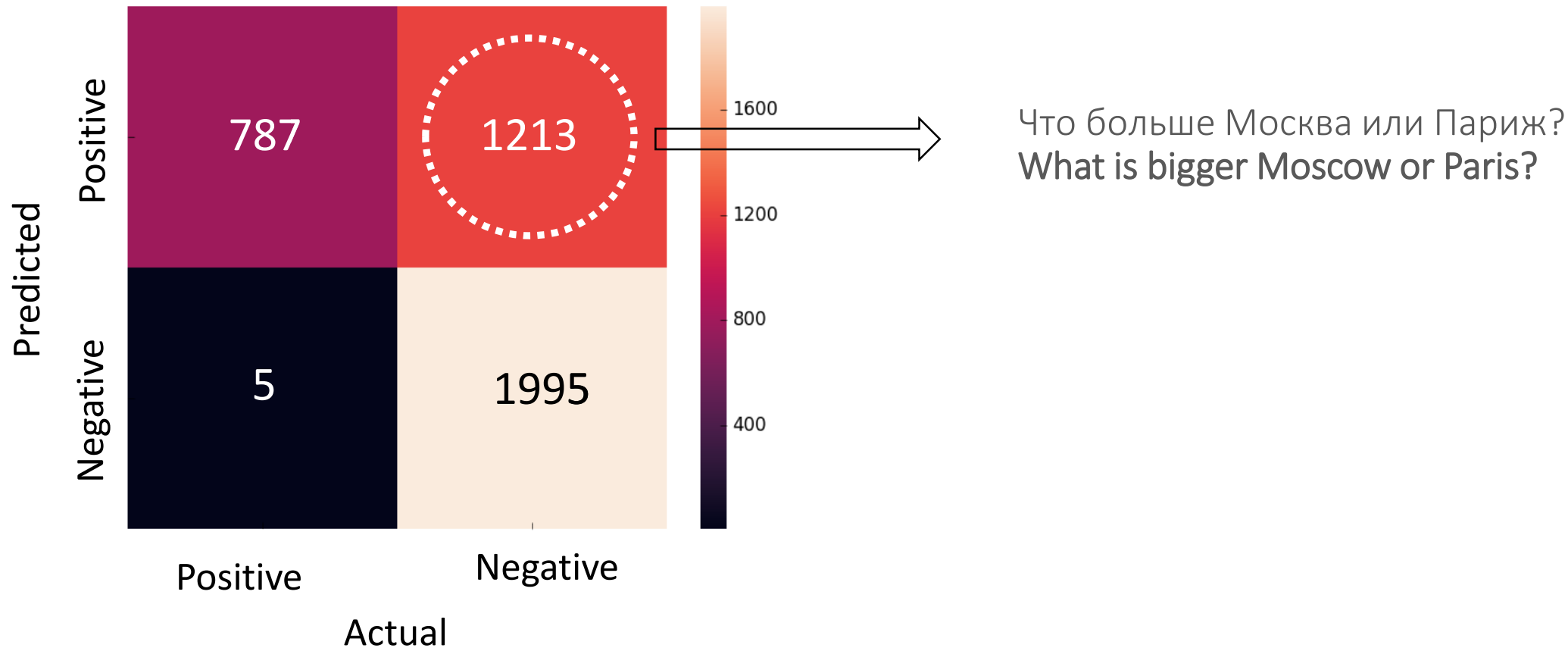
1. Input: a question
2. If it matches any of the patterns, mark it as comparative and non-comparative otherwise



# Pattern-based classification of reasoning comparative questions



# Pattern-based classification of reasoning comparative questions



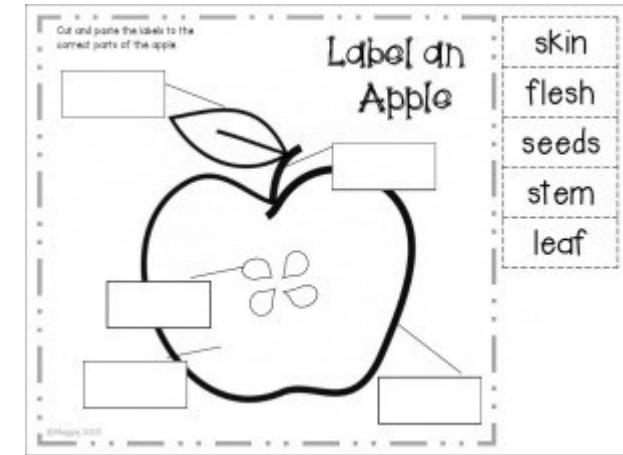
Solution – using a supervised machine learning



# Dataset preparation for labeling

Dataset was built combining:

- Random sampling
- Pattern matching
- Pattern-free mining



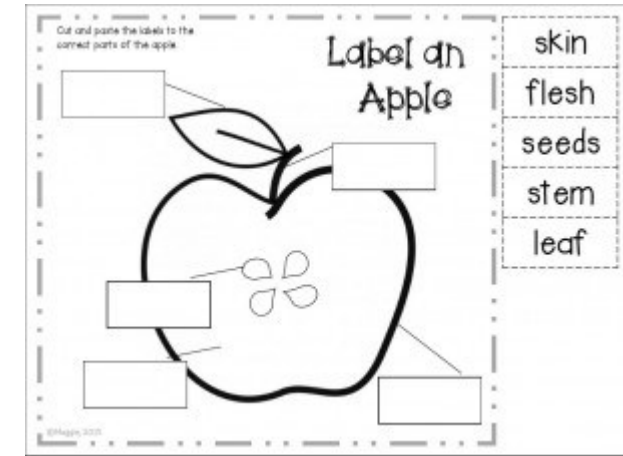
# Dataset preparation for labeling

Dataset was built combining:

- Random sampling
- Pattern matching
- Pattern-free mining

Yandex: **1%** comparative

Otvety: **3%** comparative



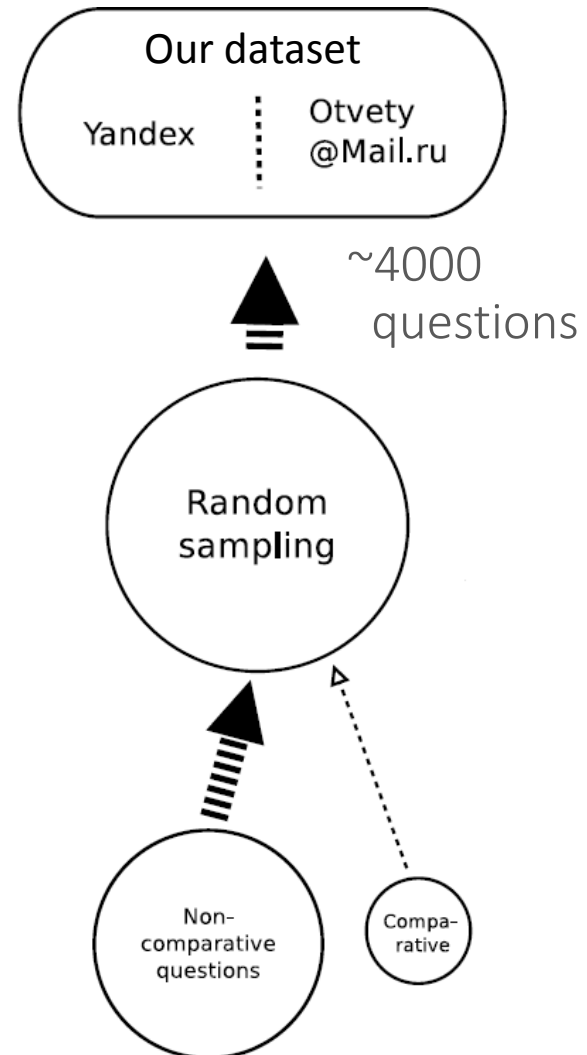
Human experts labeled data manually into classes: comparative, non-comparative and not clear.

10,000 examples have been labeled; 65% negative, 35% positive

# Dataset preparation for labeling

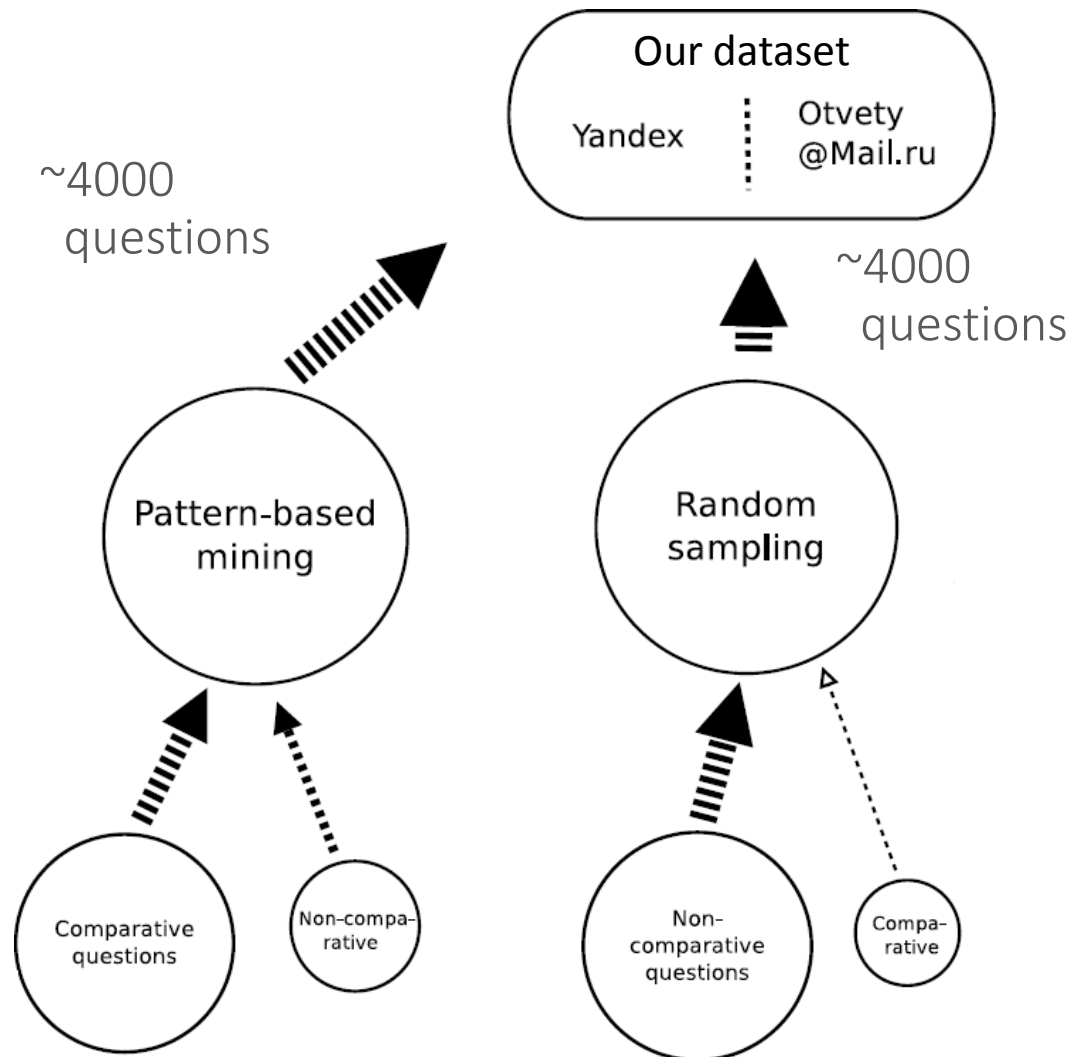


# Dataset preparation for labeling

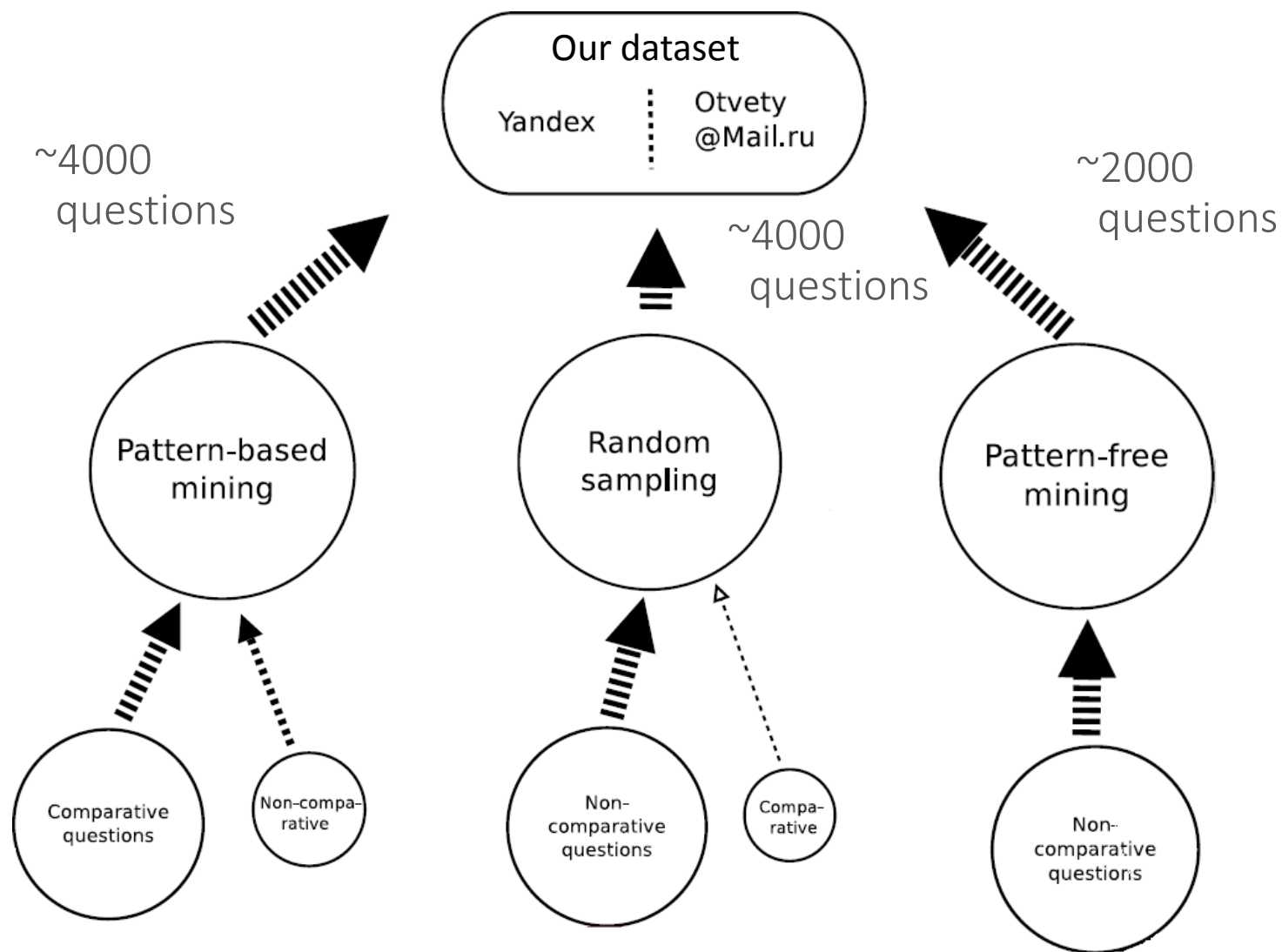




# Dataset preparation for labeling



# Dataset preparation for labeling



# Data annotation for the machine learning-based classification

Features:

Bag-of-words and part-of-speech tags:  
words are lowercased, normalized, transliterated

# Data annotation for the machine learning-based classification

Features:

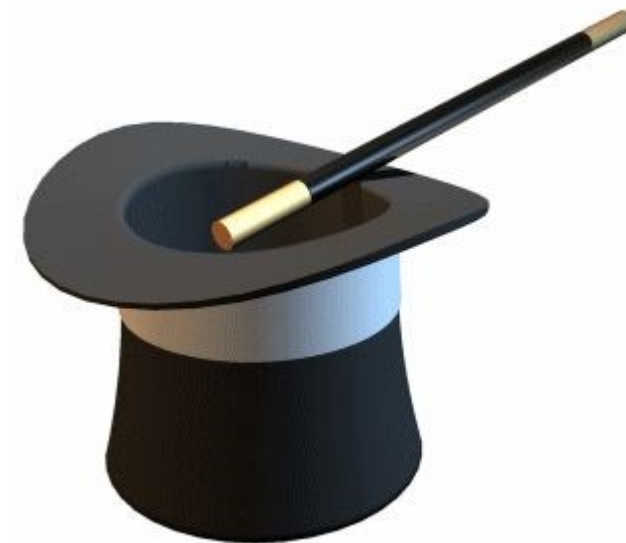
Bag-of-words and part-of-speech tags:  
words are lowercased, normalized, transliterated

“Trick”:

~~All adjectives and adverbs in comparative form~~



POS tag



# Performance of the linear SVM trained with different features

Measure,%	POS tags	Words, POS tags, unigrams, w/out comp. subst.	Words, POS tags, unigrams, w/ comp. subst.	Words, POS tags, uni-, bi-, trigrams, with comp. subst.
Accuracy	79	92	93	93
Precision	76	91	92	92
Recall	74	91	92	92
F1	79	91	92	92

# Performance of the linear SVM trained with different features

Measure,%	POS tags	Words, POS tags, unigrams, w/out comp. subst.	Words, POS tags, unigrams, w/ comp. subst.	Words, POS tags, uni-, bi-, trigrams, with comp. subst.
Accuracy	79	92	93	93
Precision	76	91	92	92
Recall	74	91	92	92
F1	79	91	92	92

# Machine learning-based classification of the labeled dataset

Terms: 22,000

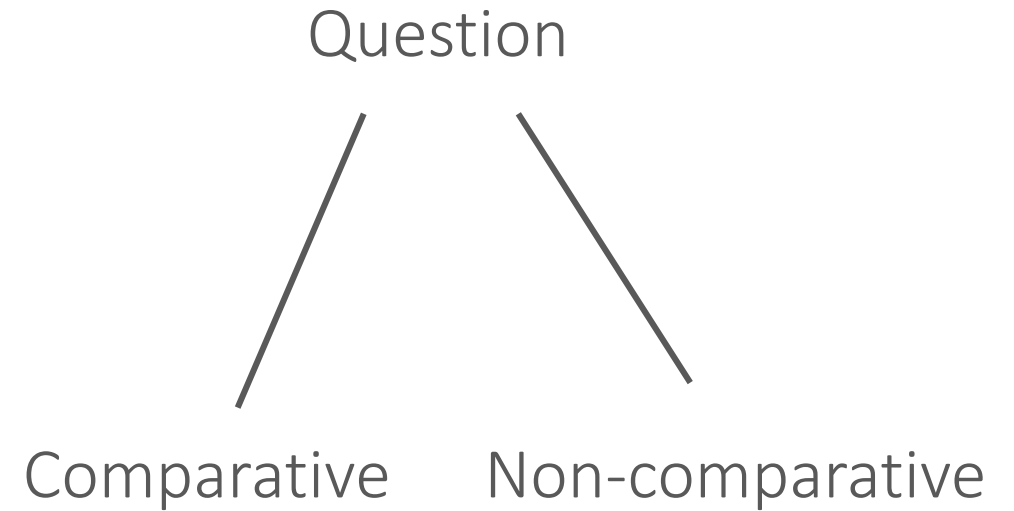
10,000 examples have been labeled; 65% negative, 35% positive

5-fold cross validation



Measure, %	Linear SVM	Logistic regress.	Bernoul. NB
Accuracy	93	92	93
Precision	92	92	92
Recall	92	92	91
F1	92	92	92

We use our pre-trained model to classify questions in our corpora





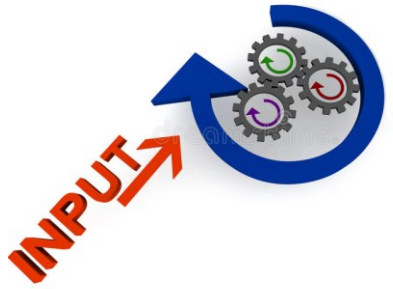
# ML-based classification of direct reasoning comparative questions



1.5 billion question queries to the Russian search engine

7.7 million questions from the Russian CQA

# ML-based classification of direct reasoning comparative questions



1.5 billion question queries to the Russian search engine  
7.7 million questions from the Russian CQA

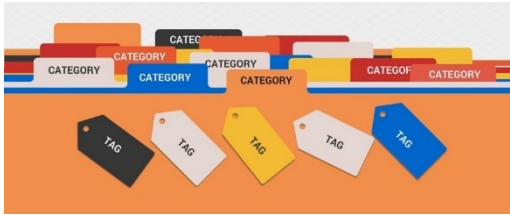


14 topic-like categories following [Völske et al., 2015]

# ML-based classification of direct reasoning comparative questions



1.5 billion question queries to the Russian search engine  
7.7 million questions from the Russian CQA



14 topic-like categories following [Völske et al., 2015]



Pre-trained supervised machine learning-based classifier

# Comparative questions on the Russian web



Search engine

Category	Comp. Q's, %
Education	3.2
Society & Culture	2.6
Consumer electronics	2.4



Community Q & A

Category	Comp. Q's, %
Consumer electronics	6.1
Cars & Transportation	5.4
Home & Garden	4.6

# Comparative questions on the Russian web

**Яндекс**

Search engine

Comparative questions 14 million		Non-comparative q. 1.5 billion	
Q-word	Ratio, %	Q-word	Ratio, %
What	33	How	50
How	23	What	11
Which	14	Which	7

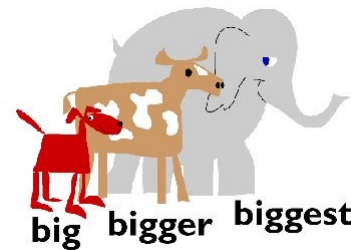
# Comparative questions on the Russian web

**Яндекс**

Search engine

Comparative questions 14 million		Non-comparative q. 1.5 billion	
Q-word	Ratio, %	Q-word	Ratio, %
What	33	How	50
How	23	What	11
Which	14	Which	7

The majority of the comparative questions start with **what**



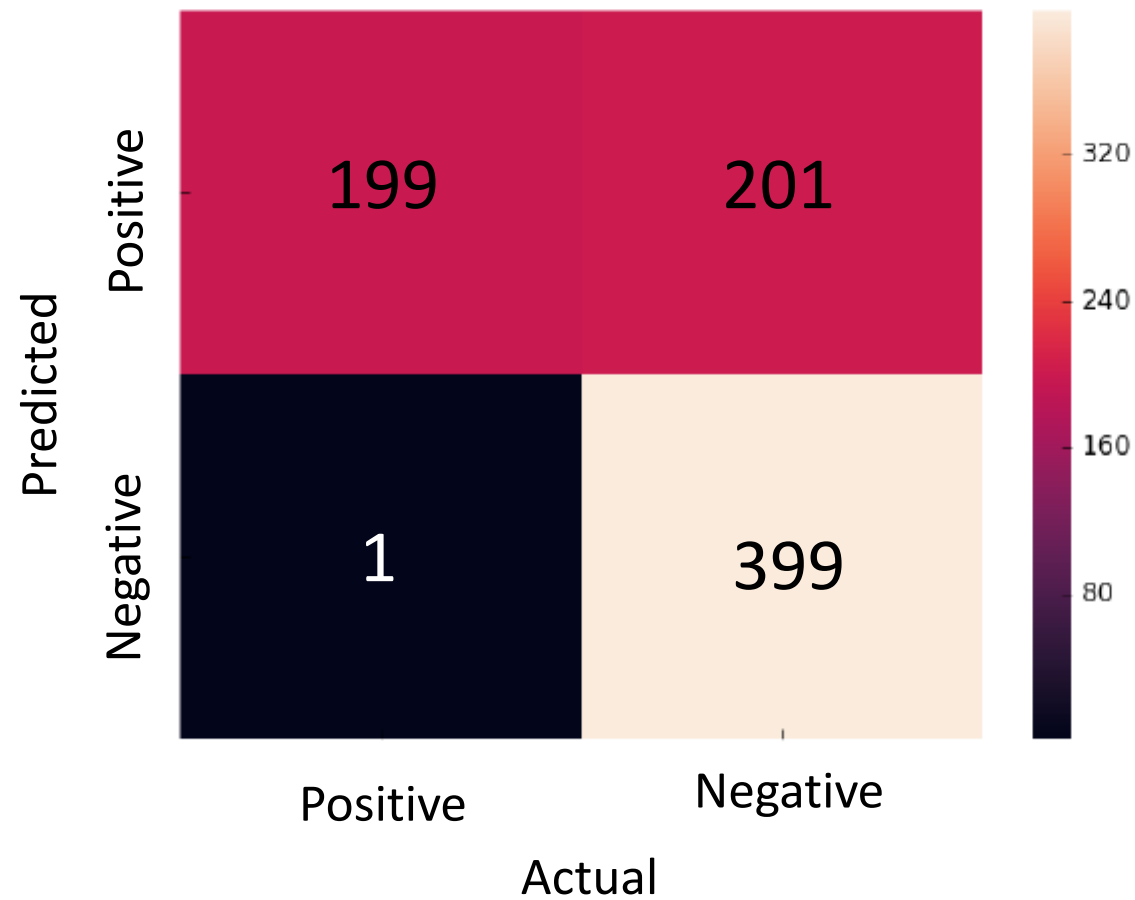
The most non-comparative are **how/how to**-questions

# Comparative question ML-based classifier performance

**Яндекс**

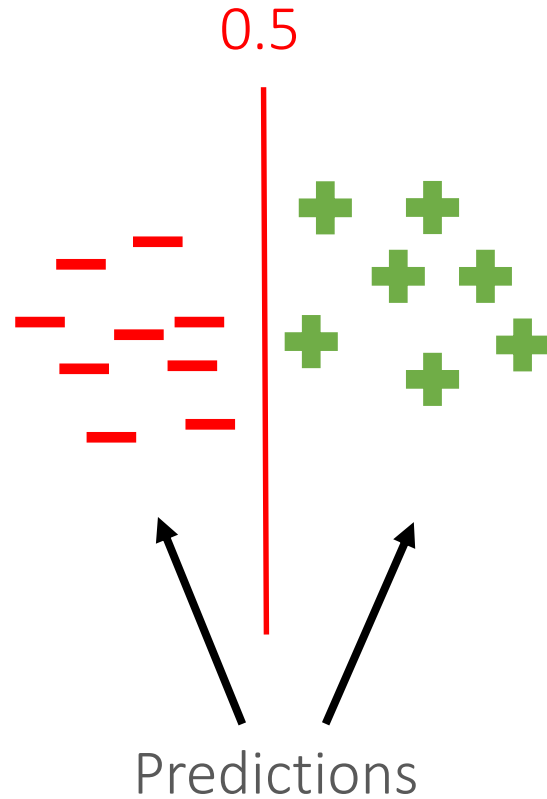
Search engine

Measure	Value, %
Accuracy	75
Precision	55
Recall	99
F1	65



# Reducing the false-positive predictions

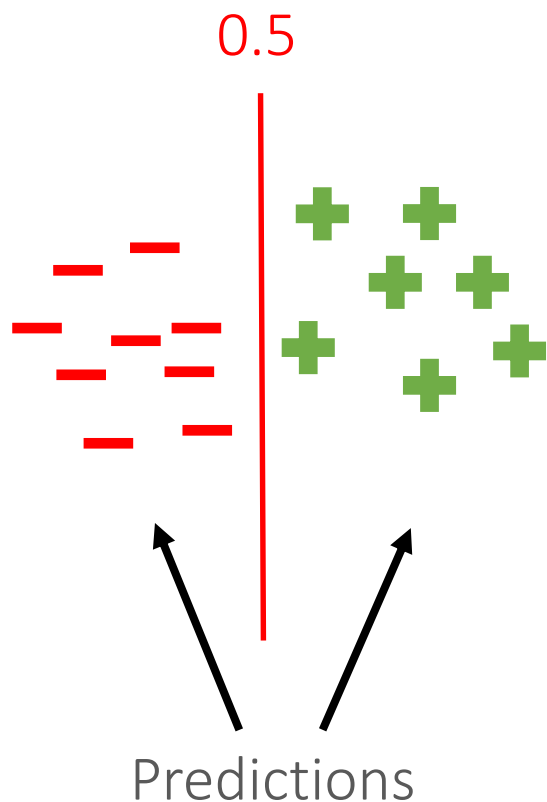
Classifier confidence score



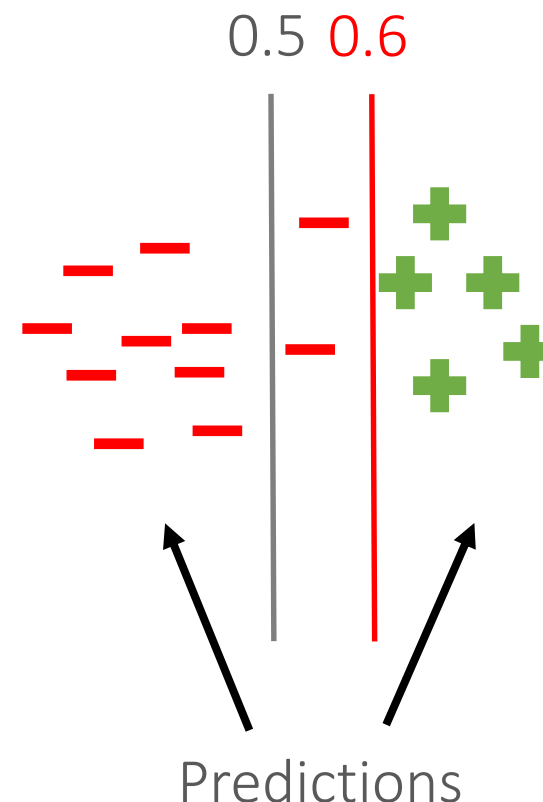


# Reducing the false-positive predictions

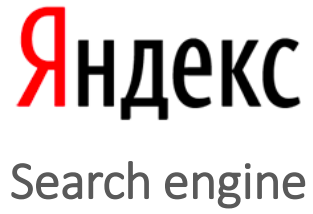
Classifier confidence score



Classifier confidence score



# Comparative questions on the Russian web



1.5 billion question queries

1% – reasoning direct comparative questions

1 comparative question / 2 sec.

# Comparative questions on the Russian web

**Яндекс**

Search engine

2018 >> 2012

1.5 billion question queries

1% – reasoning direct comparative questions

1 comparative question / 2 sec.

## Comparative questions on the Russian web

- ✓ Pattern-based mining / comparative patterns
- ✓ Training dataset manual labeling
- ✓ Machine learning for comparative question classification
- ✓ Qualitative analysis: comparative questions on the Russian web

# Comparative questions on the Russian web

- ✓ Pattern-based mining / comparative patterns
- ✓ Training dataset manual labeling
- ✓ Machine learning for comparative question classification
- ✓ Qualitative analysis: comparative questions on the Russian web



*thank you!*

# Comparative Questions/Question Queries

Backup

# Comparative Questions/Question Queries

**Table 3.1:** Pattern-based classifier performance evaluation on the Yandex corpus.

Performance measures, %	
Measurement	Value
Accuracy	72
Precision	44
Recall	100
F1 score	61

# Comparative Questions/Question Queries

## Machine learning-based classification

### Feature engineering:

**Table 4.1:** Performance of the linear SVM trained with different feature sets.

Measure	Performance measures			
	Words, POS tags, subst. comp, unigrams	POS tags only	Words, POS tags, without subst. comp, unigrams	Words, POS tags, subst. comp, uni-, bi-trigrams
Accuracy	0.9259	0.7928	0.9223	<b>0.9261</b>
Precision	0.9176	0.7556	0.9144	<b>0.9187</b>
Recall	<b>0.9194</b>	0.7436	0.9145	0.9192
F1 score	0.9180	0.7870	0.9142	<b>0.9183</b>



# Comparative Questions/Question Queries

## Machine learning-based classification

**Table 4.2:** Comparison of the classifiers' performance on the labeled dataset.

Measure	Performance measures					
	SVM Linear	SVM RBF	Decision Tree	Bernoul. NB	Logistic Regres.	Grad. Boost
Accuracy	<b>0.9259</b>	0.9197	0.9051	0.9008	0.9241	0.9250
Precision	0.9176	0.9116	0.8953	0.9052	0.9165	<b>0.9207</b>
Recall	0.9194	<b>0.9117</b>	0.8982	0.8744	0.9165	0.9131
F1 score	<b>0.9180</b>	0.9115	0.8950	0.8867	0.9164	0.9167

# Comparative Questions/Question Queries

**Table 4.3:** Comparative questions in the Yandex corpus sorted by categories.

Yandex: 1,500,825,102 questions.			
Category	Questions in category	Comparative questions	Comparative ratio, %
education	98,450,656	3,134,609	3.2
society_culture	97,443,356	2,563,935	2.6
consumer_electronics	99,363,186	2,414,295	2.4
home_garden	170,496,347	4,056,252	2.3
health	129,091,266	2,654,259	2.1
family_relationships	72,357,503	1,467,107	2.0
cars_transportation	146,408,887	2,829,809	1.9
entertainment_music	94,118,056	1,603,967	1.7
sports	43,613,556	695,861	1.6
business_finance	137,092,094	2,104,212	1.5
beauty_style	98,075,280	1,312,472	1.3
adult	55,936,554	717,097	1.3
computers_internet	145,029,127	1,162,806	0.8
games_recreation	113,349,234	814,506	0.7
Total/Average		27,531,187	1.8

# Comparative Questions/Question Queries

**Table 4.5:** Comparative questions in the Otvety@Mail.ru corpus sorted by categories.

Otvety@Mail.ru: 7,671,254 questions.			
Category	Questions in category	Comparative questions	Comparative ratio, %
consumer_electronics	280,688	17,133	6.1
cars_transportation	395,263	21,522	5.4
home_garden	565,664	25,766	4.6
sports	278,926	12,067	4.3
society_culture	943,082	39,796	4.2
education	363,226	14,744	4.1
adult	981,292	39,062	4.0
beauty_style	311,010	11,324	3.6
business_finance	409,096	13,580	3.3
health	419,388	13,039	3.1
family_relationships	1,250,761	37,987	3.0
entertainment_music	704,830	19,385	2.8
computers_internet	476,390	10,245	2.2
games_recreation	291,638	4625	1.6
Total/Average		280,275	3.7

# Comparative Questions/Question Queries

## Choosing a classifier for the corpora

We randomly sample 100 positive and 100 negative predictions and calculate performance. Even though the **support vector machine** demonstrated a very high recall of 100%, its overall performance was rather poor with 33% F1 measure. **Logistic regression**, in contrast, has shown much better performance as 85% F1 measure.

# Comparative Questions/Question Queries

## Categories examples

### Education:

в чем разница между восстанавливающий и восстановительный?  
**what is the difference between recreational and reconstructive?**

чем мед колледж отличается от мед академии?  
**how a medical college differs from a medical academy?**

### Society & culture:

у кого лучше развито невербальное общение у мужчин, женщин или детей?  
**who has a better developed nonverbal communication: men, women, or children?**

в чем различие между японской и китайской кухней?  
**what is the difference between Japanese and Chinese cuisine?**

# Comparative Questions/Question Queries

## Performance measures

Accuracy refers to how close overall predicted classes to actual or how accurately the classifier can distinguish between comparative and non-comparative questions in a binary classification task. Accuracy is a good measure only if the amounts of false positive and false negative predictions are almost same. It is calculated as

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

Precision refers to how close to each other measurements for each class or, in other words, how good the classifier performs in predicting comparative and non-comparative questions. Low precision usually indicates a high number of the false positive predictions. It is defined as

$$Precision = \frac{TP}{TP + FP}$$

Recall describes how many of the truly comparative questions our classifier labeled. Low recall indicates many false negative predictions. It is calculated as

$$Recall = \frac{TP}{TP + FN}$$

F1 score refers to the harmonic mean between the precision and the recall and is useful to use when classes are distributed unevenly. It is calculated as

$$F1 \text{ score} = 2 \times \frac{recall \times precision}{recall + precision}$$

Table 3.1 demonstrates our pattern-based classifier performance on the whole Yandex corpus based on 1000 randomly sampled and manually checked classification results.