Automatic Answering of Students' Questions by Using an Archive of Questions

Author: **Adrián Huňa**
Supervisor: **Ivan Srba**

Motivation

- Students' questions posted online
- Large online communities
  - MOOCs, forums, Askalot
- Teacher overload
  - Questions repeat in time
- Valuable knowledge in archives could be utilized better
- Automatically detect similar questions and provide answers from archive
  - **Utilize specifics of educational domain** (teachers, time similarity, ...)

Proposed method

Question or discussion?

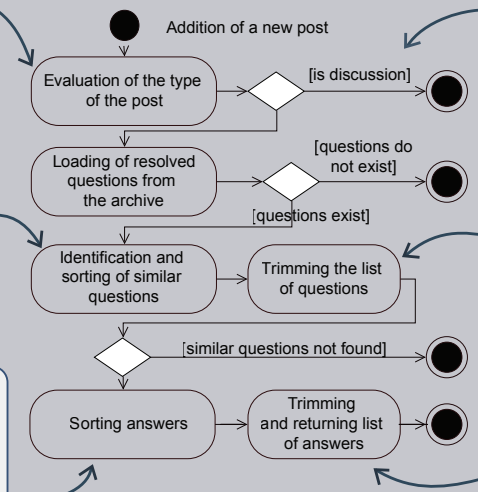
Information provided by user. (Could use classification.)

Classification

- Any classifier can be used
  - Random Forest, SVM, Naive Bayes
- 91 features
  - Textual similarity (TF-IDF, Glove)
  - Part of speech (nouns, verbs)
  - Bigrams
  - Categorical similarity
  - Time similarity
  - Question's title, text, "asking part"
- Feature selection
  - GBM, Recursive feature elimination

Learning to rank problem

- SVMrank
- Same features as for similar questions classification and new features such as: answer votes, answerer is teacher, answer length, etc.
- 81 features - feature selection again



We try to answer only questions which have an objectively true answer.

Only questions with some probability of similarity must be kept.

Return as many answers as suitable for a given use case.

Experiment

- Scraped data from a MOOC provider
- Offline experiment on 2 iterations of a Computer Science course
  - **1284** users asked or answered **2255** posts - **1612** were questions. The posts were replied by **3599** answers.
- Question similarity classification and ranking evaluated separately
- Manual data labeling
  - Post type (question/discussion)
  - Question similarity
    - For top 12 TFIDF similar pairs
    - Some data labeled by 2 annotators
  - Answer helpfulness
- Evaluated only on labeled data

Similar questions classification

- 78 similar and 218 not similar questions in test)
- Most important features are:
  Text+title TFIDF similarity, Noun intersection in text to unique words ratio, Title similarity based on Glove, Noun intersection in title to unique words ratio, Text TFIDF similarity

	Correcly answered	Unanswered	Incorrectly answered	Success@1	Precision for similar	Recall for similar
Random Forest (RF)	13	1	1	0.7789	0.8666	0.1666
RF without educational	13	2	2	0.7721	0.7647	0.1666
SVM	15	1	2	0.7857	0.8333	0.1923
SVM without educational	15	0	4	0.7789	0.7894	0.1923

Answer ranking

- 58 comparisons in test
- Features used after feature selection (sorted by importance):
  Author's best answer ratio, Answer length, Time since original question was asked, Answerer is teacher, Answer votes, Question marks count

	P@1	nDCG
With education features	0.7758	0.9007
Without education features	0.7413	0.8846

Conclusion

- Online student communities are different than communities on the open web
- Educational features improve performance

Future work

- Utilize answer text for similar questions classification
- Use Glove based similarity when sorting answers
- Evaluation of data from different domain
- Online experiment

STU FIIT

PeWe@FIIT
personalized web group