Project Plan Presentation

Jingheng Pan
jingheng.pan@studium.uni-hamburg.de

Adrian Lindloff
adrian.lindloff@studium.uni-hamburg.de

Lijunnan Bai
lijunnan.bai@studium.uni-hamburg.de

Structure

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

1. Motivation & Problem Statement
2. Demonstration of the original Storyfinder Plugin
3. Use-Case
4. Architecture
5. Functionality
6. Project Outline
7. Future Work

What is the “Adaptive Storyfinder” Project?

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Personalized content recommendation system to serve the individual interests of users by finding “stories”/ articles based on reading habits

  → we focus on re-ranking these articles

  → enhance user engagement and satisfaction

- Original Storyfinder is “only” an archive of visited websites

What is the “Adaptive Storyfinder” Project?

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Personalized content recommendation system to serve the individual interests of users by finding “stories”/ articles based on reading habits

  → we focus on re-ranking these articles

  → enhance user engagement and satisfaction

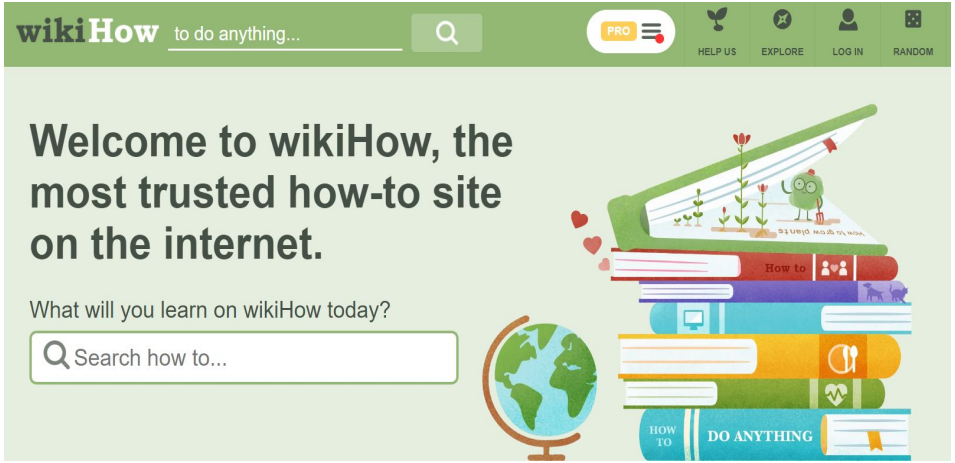
- Original Storyfinder is “only” an archive of visited websites
- “Adaptive” Storyfinder shall incorporate human behavioral data and preferences

WikiHow

- Link: https://www.wikihow.com/Main-Page
- Website featuring how-to articles on a variety of topics

- Practical to gather information regarding all kinds of interests

- Contains large amounts of data → existing datasets



Why is this project useful?

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Use-Case of a typical WikiHow search:

- Person wants to know more information about a specific topic
- Has to spend some time searching for a suitable and enjoyable article

=> Goal of the Adaptive Storyfinder:
- User preference-oriented ranking of results
- (Additionally display more recommended content)

→ overall more captivating and enjoyable experience

Architecture

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Backend:

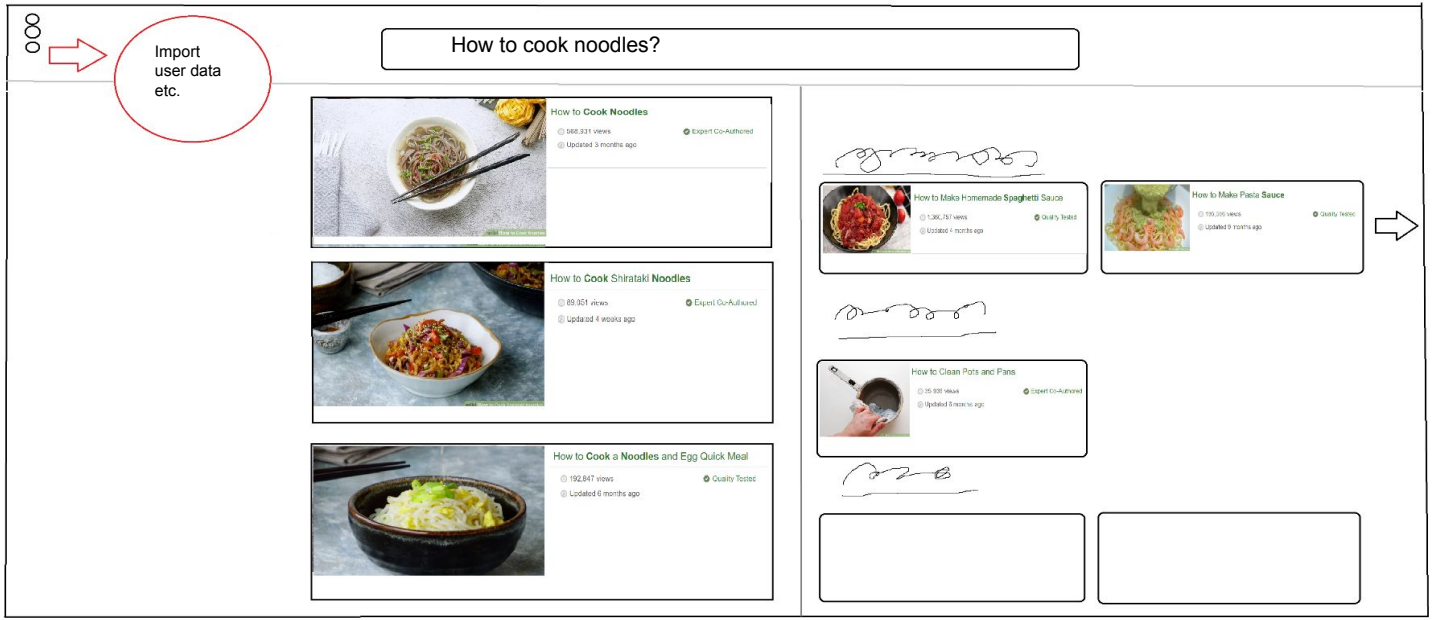
- Django (Python web framework)
- Docker
- PyTorch for ML

Frontend:

- Vue.js and some toolkits depending on the situation

Functionality

- Re-rank given result
- Recommendations



Re-rank given results

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Re-rank goal: Article with the highest similarity to the preferences is at the top

- Create scores without ML
  - Create scores for predefined features of each presented article (f.E. amount of images etc.) by leveraging a large WikiHow dataset
  - Create user preferences
- Create user preference with ML
  - Create wikihow article embeddings and user embeddings with ML

Milestones

- Workable environment with WikiHow Api

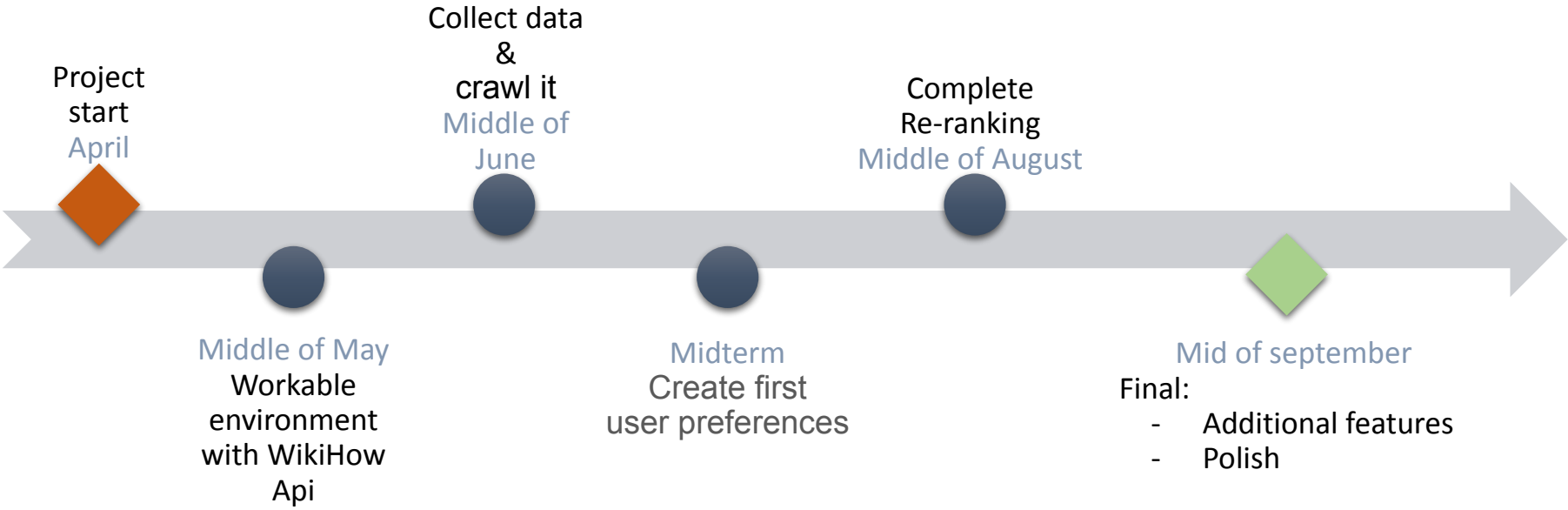
Create user preferences:

- Collect individual user data with Plugin → duration spends on websites
- Collect the website features (f.E. amount of images) → crawler
- Build user preferences

Re-rank results:

- Calculate similarity between user preferences & result websites

Project Outline



Future Work

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

More features/ thoughts:

- Make a custom reading page // Summarize an article based on preferences
- Recommendation part: actively generate new article suggestions

- Implement user/ data privacy or at least inform the users

  about the scope of data usage

- Include other sources of information beyond the use of WikiHow

Questions

Recommendations

- Collaborative Filtering(CF)
  - UserCF
  - ItemCF

	item1	item2	item3	item4	item5
u1	5	3	4	4	??
u2	3	1	2	3	3
u3	4	3	4	3	5
u4	3	3	1	5	4
u5	1	5	5	2	1

- Matrix Factorization(MF)
  - Matrix decompose to implicit vector
  - General way: Machine learning from training dataset

	x1	x2	x3	x4
u1	1	2	3	4

	x1	x2	x3	x4
Item1	4	3	0	0