



evalXAI

AI Bench-marking Platform

Group Members:

Vincent Beyer
Mochtaba Aljorani
Konrad Goldammer
Mateusz Plikus
Nico Pelleriti

Supervisors:

Rick Wilming
Benedict Clark

Summer Term 2024

Technical University of Berlin

evalXAI Project Journal

Week 17.04.-23.04. (first session)

- Got comfortable with the group
- Helped setting up a discord server
- Met on discord and determined a distribution of roles

Week 24.04.-30.04.

- Cloned the exact GitHub repo and installed all the dependencies
- Looked at the documentation of the previous group and created working Docker images for the backend, frontend, and worker
- Tried to understand the code base and the whole structure of the work from the previous group

Week 01.05.-07.05.

- Did extensive research on Django because I had never used it before (YouTube tutorials, official documentation, etc.)

Week 08.05.-14.05.

- Did further research on Django and other technologies that are used in the provided code base
- Fixed API calls in the Frontend by adding CORS middleware in the Django backend

Week 14.05.-21.05.

- Added a user API, which is responsible for the user management in the backend (registering, logging in, logging out, getting user information, etc.)

Week 22.05.-28.05.

- Implemented the Mantine component library (<https://mantine.dev/>) in the React frontend
- Refactored login and register pages, which Mochtaba created, by using Mantine components (among other things)
- Created video for presentation of results for milestone I

Week 29.05.-04.06. (milestone I)

- Fixed the session handling in the Frontend
- Implemented a competition creation page, which uses the competition creation route in the backend

Week 05.06.-11.06.

- Added a user context in the Frontend, which manages app level state regarding user metadata and prevents duplicate API calls

Week 12.06.-18.06.

- Refactored existing challenges overview page
- Added mock data for testing purposes
- Came up with outline for challenges detail page

Week 19.06.-25.06.

- Added grid for displaying all existing challenges with new card component in the challenges overview page
- Refactored the home page, header, and footer, to make it look nicer
- Implement the planned structure for the challenges details page (main section with description and thumbnail, „Metadata“ section, „Your Submission“ section, „Leaderboard“ section)

Week 26.06.-02.07. (milestone II)

- Added modal for XAI methods inside the „Your Submissions“ section of the challenge details page

Week 03.07.-09.07.

- Fixed the view for submitting XAI methods and getting the computed score in the Django backend
- Refactor the add submissions modal by adding details like the loading bar
- Fixed the dynamic leaderboard in the challenge detail page (it now contains actual data that is stored in the „Scores“ table from the data base)

Week 10.07.-16.07.

- Improved the entire XAI method submission work flow by adding minor details in the frontend
- Fixed weird problem concerning the „invalid CSRF token“ error with API calls, so that making API calls from the Frontend while being logged is now working

Week 17.07.-23.07. (final presentation & poster session)

- Worked on poster for the poster presentation (provided contents for the sections: architecture diagram, example usage, results)

evalXAI Project Journal - Mochtaba Aljorani

Weeks 17.04. - 14.05.

Preparatory Work:

- Task allocation and initial team introduction (first Discord meeting).
- Cloned the repository and set up the project environment, including all dependencies. Addressed local setup issues, such as those with WSL.
- Conducted preliminary research on Docker, Django, and JavaScript to better understand their application in our project.
- Examined the frontend code and tried to understand how the platform works

Week 15.05. - 21.05.:

- Developed the initial versions of the login and registration pages.
- Connected it to the Backend.

Week 22.05. - 28.05.:

- Addressed frontend issues, such as duplicate headers, and began development on the Profile Page.

Week 29.05. - 04.06. (Milestone I):

- Completed the Profile Page and included a placeholder for a future profile picture.

Week 05.06. - 11.06.:

- Retrieved user data from the User Context and displayed it on the profile page.
- Attempted to implement user profile picture uploads and other user data enhancements, which were not successful due to backend priorities.

Weeks 12.06. - 25.06.:

- Implemented data fetching logic for the challenge cards and displayed this data on the challenge page.

Week 26.06. - 16.07. (Milestone II):

- Implemented an initial version of the data table using mock data.

Week 17.07. - 23.07. (Final Presentation & Poster Session):

- Assisted in creating the project poster and wrote sections detailing the methods used and the current state of the project.

evalXAI Journal: Vincent Beyer

Role: Backend/Worker

Week 17.04.-23.04.

- Got to know the group
- Set up a discord server for meetings
- Distributed roles

Week 24.04.-30.04.

- Cloned the exact GitHub repo and installed the dependencies
- Tried to understand provided code
- Read the XAI-tris research paper

Week 01.05.-07.05.

- Learned Django and Docker

Week 08.05.-14.05.

- Learned more Django and Docker

Week 14.05.-21.05.

- Learned even more Django and Docker
- Helped Matthew with Django and the database

Week 22.05.-28.05.

- Started to implement multiple workers (Idea: spawning multiple worker containers from one backend container)
- Edited Dockerfiles

Week 29.05.-04.06.

Encountered problem: "Not supported URL scheme http+docker"

Week 05.06.-11.06.

- Found solution: upgrade docker api and requests versions
- Implemented logging in my spawn worker function

Week 12.06.-18.06.

- Finished implementing spawning worker containers from the backend container

Week 19.06.-25.06.

- Fixed error/warning messages when creating and running worker containers

Week 26.06.-02.07.

- Finished, added comments and pushed multiple worker functionality

Week 05.06.-11.06.

- Added a worker limit
- Problem: getting the score from worker container to backend container
 1. Solution (not final): docker network between backend and worker containers to allow worker container to make api call with score (see "dev-worker" branch in GitHub repo)
 2. (final) Solution from Konrad: get logs from worker container and search for score

Week 12.06.-18.06.

- wrote text for poster ("about.info" section)

After: adjusted readme, comments in Dockerfiles/functions and deleted unused functions

Week 17.04.-23.04.

- Got comfortable with the group
- Role distribution

Week 24.04.-30.04.

- Project setup and analysis
- First testing and docker setup on windows

Week 01.05.-07.05.

- Researching the Xai-tris paper
- Understanding the topic of XAI

Week 08.05.-14.05.

- Django YouTube tutorials research
- Http requests and client-server communication research for better understanding

Week 14.05.-21.05.

- Researching ways of file storage compatible with Docker

Week 22.05.-28.05.

- Implemented Challenge Management Features with amazon S3 cloud storage
- Endpoints for fetching and adding challenges, formular page for submitting new challenges

Week 29.05.-04.06. (milestone I)

- Fixed a minor spelling bug

Week 05.06.-11.06.

- New backend endpoint for fetching all challenges as JSON

Week 12.06.-18.06.

- Researching docker and django local storage techniques because of the need to change the amazon cloud storage into local storage

Week 19.06.-25.06.

- Changed the storage from amazon s3 cloud storage to local file storage

Week 26.06.-02.07. (milestone II)

- Score management endpoints added

Week 03.07.-09.07.

- Default file bug fix
- urls.py score endpoints update

Week 10.07.-16.07.

- updated the README.md in the backend for future users

Week 17.07.-23.07. (final presentation & poster session)

- Worked on the poster design

Week 17.04.-23.04.

- First meeting, overview, introduction and forming of group.
- Distributed roles according to interests and knowledge.

Week 24.04.-30.04.

- Setup of Project, including cloning the repo and running the project (OS: Ubuntu).
- Read existing code and documentation.

Week 01.05.-07.05.

- Read the paper "XAI-TRIS: Non-linear image benchmarks to quantify false positive post-hoc attribution of feature importance".
- Learned about the workflow with the given jupyter notebook.

Week 08.05.-14.05.

- Read the paper "Scrutinizing XAI using linear ground-truth data with suppressor variables"
- Learned the basics of the torchhub api.

Week 14.05.-21.05.

- Started implementing torchhub api features to enable model loading via the torch module.
- Learned about the current training procedure.

Week 22.05.-28.05.

- Implemented model loading for the remaining architectures and set up tests to confirm the correct models are loaded.
- Added documentation on how to load the models.

Week 29.05.-04.06.

- Began adjusting the training procedure.
- Read about basic XAI methods.

Week 05.06.-11.06.

- Tested currently implemented XAI method LRP with the given models.

Week 12.06.-18.06.

- Read LIME and SHAP papers.
- Learned about Captum and ways to implement standard XAI methods.

Week 19.06.-25.06.

- Added documentation.

Week 26.06.-02.07.

- Cleaned up code, updated training procedure of models.
- Changed the score function to accept different models and datasets.

Week 05.07.-11.07.

- Added other XAI metrics, IMA and Precision from Benny's paper.
- Implemented tests for these metrics.

Week 12.07.-18.07.

- Added more functionalities to the score method. Submitted code is now executed on the server. (MAJOR security risk!!!)
- Added documentation.
- Worked on text for ML part of our project.

