



How to organize our work?



MI²DataLab Winter Seminar 2022



Rules for today

1. One statement == one treat.
2. Focus on currently discussed topic (there will be many).
3. No blaming, we want to reflect and draw conclusions!
4. Everything is interactive, so add your own rule if you want ;)

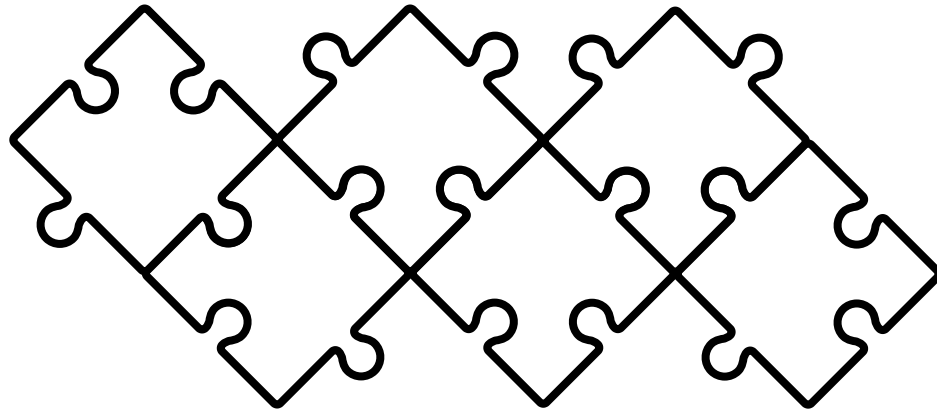
What did we do right?

What worked in our plan?

What made these things successful in implementation?

Pytanie o Recenzjach
BANIECKI HUBERT - Hackathon w Włocławku / Październik
BIECEK PRZEMEK - Polityka 22 2020, 2021
DOMITRZ WITALIS - PRZYGOTOWANIE NARZĘDZIA DO
GIZIŃSKI STANISŁAW - Baza wiedzy
GRZYB MATEUSZ - karta z aktywnościami, INTEGRACJA
KOZAK ANNA - BAZA WIEDZY (REKRYWALNE MAT.)
KOZŁOWSKI ADAM - BAZA WIEDZY - INSTRUKCJE TECH
KRZYŻYŃSKI MATEUSZ - WELCOME BOOK (REKRYWALNA
- HACKATHON - BLOG AKTYWIZACJA)
ŁANIEWSKI STANISŁAW - Party Planning Committee
LY THIEN
RUCZYŃSKI HUBERT - Party Planning Committee
SKIBICKA WERONIKA - KALENDARZ
SOBIESKI BARTEK - Seminarium (org + coord.)
SPYTEK MIKOŁAJ - BAZA WIEDZY - INSTRUKCJE TECH
STĄCZEK MATEUSZ - HACKATHON
STAŃDO ADRIAN
WILCZYŃSKI PIOTR @H12.171
WIŚNIOŚ EMILIA - Party Planning Committee
WOŹNICA KATARZYNA - seminarium (org + coord.)
ZÓŁKOWSKI ARTUR - NARZĘDZIE BAZY WIE

How do we function now?



Website

- @Hubert Baniecki is a policeman of website and tries to update it periodically

TO DO:

- tab with subjects of diploma thesis
- updating Solution tab - list of cooperation and projects

Roadmap

- Who should have access to this repo? -> if somebody is interested in research in lab (contact with @Przemyslaw Biecek)
- For newcomers: when somebody is added to slack, somebody introduces him to team and also cares about this person

The screenshot shows a GitHub repository page for 'MI2DataLab / roadmap'. The repository is private and has 4 watchers, 0 forks, and 0 stars. The navigation bar includes links for Code, Issues (31), Pull requests, Actions, Projects (1), Security, and Insights. The main content area is titled 'Papers & Theses & Books' and displays a Kanban board with several project cards. Each card represents a task or milestone, often linked to a book or paper. The cards are organized into columns: 'Initial results', 'Manuscript - draft', 'Preprint on arXiv', 'Paper submitted', and 'Paper accepted (close if published)'. Each card includes a title, a description, a number (e.g., #72, #71, #34), and the person who opened it (e.g., pbiecek, barbara3430, woznicak). Some cards have labels like 'BetaBit', 'EMA', or 'Translation'. The bottom of the image features an orange bar with the 'MI' logo in a yellow square.

MI2DataLab / roadmap Private

Watch 4 Fork 0 Star 0

< > Code Issues 31 Pull requests Actions Projects 1 Security Insights

Papers & Theses & Books
ated 13 days ago

Filter cards + Add cards Fullscreen Menu

10 Initial results + ...

- [Book] Wykresy od kuchni: how to prepare good statistical graphics #72 opened by pbiecek BetaBit
- Understanding of features determining the classification nevi vs. melanoma #69 opened by dtomaszewska

7 Manuscript - draft + ...

- [Book] ResponsibleML: Translation from R to Python #71 opened by pbiecek BetaBit EMA Translation
- [Book] Chaos game: Translation from PL to ENG #70 opened by pbiecek

1 Preprint on arXiv + ...

- T-EMDE: Sketching-based global similarity for cross-modal retrieval #34 opened by barbara3430

3 Paper submitted + ...

- Consolidated learning - a domain-specific model-free optimization strategy with examples for XGBoost and MIMIC-IV #64 opened by woznicak
- The Grammar of Interactive Explanatory Model Analysis #16 opened by pbiecek


28 Paper accepted (close if published) + ...

- [BSc] Narzędzie do analizy podobieństwa rozkładów metryk jakości algorytmów uczenia maszynowego w zadanej przestrzeni hiperparametrów #66 opened by woznicak
- Multi-omics disease module detection with an explainable Greedy Decision Forest

Article review

- it is up to date
- we should communicate this to use to papers first author remember to upload new versions

Papers with reviews

1. One directory/table row per a project/paper.
2. Directory with date (first submission) and project name (the paper's title can change). Use  !
3. Subdirectory with a consecutive version (v1, v2, ..) and venue name (e.g. v1-jmlr, v2-tjournal).
4. For reference, highlight **the author** that uploaded the paper's reviews

DATE START	PROJECT	STATUS	FUNDING	TITLE (LATEST)	AUTHORS
2022-09-19	climpol	accepted - Tackling Climate Change with Machine Learning at NeurIPS 2022	MAIR IDUB	<i>Climate Policy Radar: Pipeline for automated analysis of public climate policies</i>	Artur Żółkowski, Mateusz Krzyżiński, Piotr Wilczyński, Stanisław Giziński, Emilia Wiśnios , Bartosz Pielirski, Julian Sienkiewicz, Przemysław Biecek
2022-09-07	survshap	submitted - Knowledge-Based Systems	HOMER+xBLungs	<i>SurvSHAP(t): Time-dependent explanations of machine learning survival models</i>	Mateusz Krzyżiński , Mikołaj Spytek, Hubert Baniecki, Przemysław Biecek

Venues

- **Decision: Forget about it**

Conferences, workshops and special issues

2022

December

- 2022-12-31. SI on **Human-centred intelligent systems** in Scientific Reports <https://www.nature.com/collections/cifcfdliif>


November

- 2022-11-30. SI on **Machine learning applications in medical image analysis** in Scientific Reports <https://www.nature.com/collections/gfbjhtfjgg>

Templates of slides and posters

- Logos -> channel #ksiega-znaku
- Space to hang posters next to 317 & 316

Question: do we need communication facility? To what extent?




survex: model-agnostic explainability for survival analysis

Mikołaj Szytek¹, Mateusz Krzyżński¹, Hubert Baniecki², Przemysław Biecek^{1,2}

¹M2 AI, Warsaw University of Technology
²M2 AI, University of Wrocław


Let's talk about: explainable artificial intelligence, survival analysis, responsible machine learning



Introduction

Survival analysis is a task dealing with time-to-event prediction based on censored data. The main difference separating it from other areas of supervised learning is its output in the form of **survival probability distribution**. Survival models are predominantly used in medicine and insurance and help make critical decisions. This means that increasing trust in the models via explanations is vital, however standard post-hoc explanations cannot be applied directly due to the nature of the model's output.

survex provides model-agnostic explanations for survival models in the form of an accessible R package [1]. These are extensions of standard methods [2] adapted for models with functional output, as well as implementations of methods developed specifically for survival analysis [3, 4].

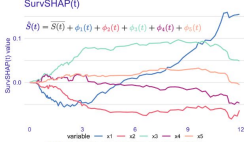


Local explanations

Local explanations help better understand model behavior around a single observation (e.g., patient).

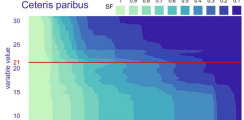
- **SurvSHAP** values show variable contributions to a model prediction at each considered time.
- **SurvLIME** explanations show local importance of variables by fitting a surrogate Cox Proportional Hazards model.
- **Ceteris paribus** plots show how the model output depends on changes of a single variable.

SurvSHAP(t)



Note: x1 has the strongest time-dependent effect

Ceteris paribus



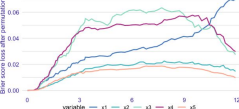
Note: Ambiguous behavior around value x4 = 24

Global explanations

Global explanations are designed to understand the general behaviour of the model for a given population.

- **Partial dependence plots** are aggregates of ceteris paribus explanations and show how changing a variable affects average model output.
- **Permutational variable importance** presents a ranking of the variables by calculating how the performance changes after permuting a variable.

Permutational variable importance






Note: Different variables rank as the most important at different timepoints

Conclusion

- **survex** incentivizes the popularization of explainability methods in domains where survival analysis is applied.
- It benefits various stakeholders e.g. physicians and biostatisticians in extracting knowledge from data and model analysis.
- In-depth analysis of the prediction helps medical personnel decide how adequate it is, in turn leading to development of **personalized medicine**.

Contact info

 miko@survex.ai@pwr.edu.pl
 <https://github.com/ModelExplanations/survex>
 www.m2.ai

References

- [1] Mikołaj Szytek, Mateusz Krzyżński, Hubert Baniecki, and Przemysław Biecek. **survex: Explainable Machine Learning in Survival Analysis**. 2022. R package version 0.1.1.
- [2] Przemysław Biecek and Tamas Bolyos. **Explanatory Model Analysis**. Chapman and Hall/CRC, New York, 2021.
- [3] Marcin Ł. Nowak, Lei Wang, and David H. Wainer. **LocalIME: A method for explaining machine learning survival models**. Knowledge-Based Systems, 203:108164, 2020.
- [4] Mateusz Krzyżński, Mikołaj Szytek, Hubert Baniecki, and Przemysław Biecek. **SurvSHAP(t): Time-dependent explanations of machine learning survival models**. arXiv preprint arXiv:2206.12590, 2022.

Acknowledgements

Special thanks to Anna Kozak, Katarzyna Woźnica and Zuzanna Kucharska for the valuable comments and discussions about this work. This work was financially supported by the NCN grant grant IP2019/00011-00/2021/01 and the NCN Sonata Bis-9 grant 2019/30/1-K/16/00012.

Code example

```
library(survex)
library(randomForestSRC)
rf_model <- rfrsrc(Surv(time, event) ~ ., data = df)
rf_explainer <- explain(rf_model)
perm_var_imp <- model_perms(rf_explainer)
plot(perm_var_imp)
```

¹ Python implementation of SurvLIME and SurvSHAP(t) methods is also available at <https://github.com/ModelExplanations/survshap>.

Wiki

- updating instruction about Eden

Home

Witalis Domitrz edited this page on Jul 13 · 15 revisions

[Edit](#)[New page](#)

Baza wiedzy

To jest strona główna wiki/bazy wiedzy labu MI².

Słowo wstępu

Krótkie wprowadzenie do obsługi GitHub wiki

Lista wszystkich stron

Po prawej stronie od tego (i każdego innego) artykułu w bazie wiedzy jest (zwinięta, lub rozwinięta) zakładka **Pages**. Zawiera ona listę wszystkich stron i artykułów. Można w niej wyszukiwać i wtedy wyszukujemy nie tylko po tytule strony, ale też po wszystkich nagłówkach na podstronach

Pages 8

Część podstron nie jest jeszcze stworzona. Linki zostały utworzone jedynie dla przykładowych podstron.

O labie

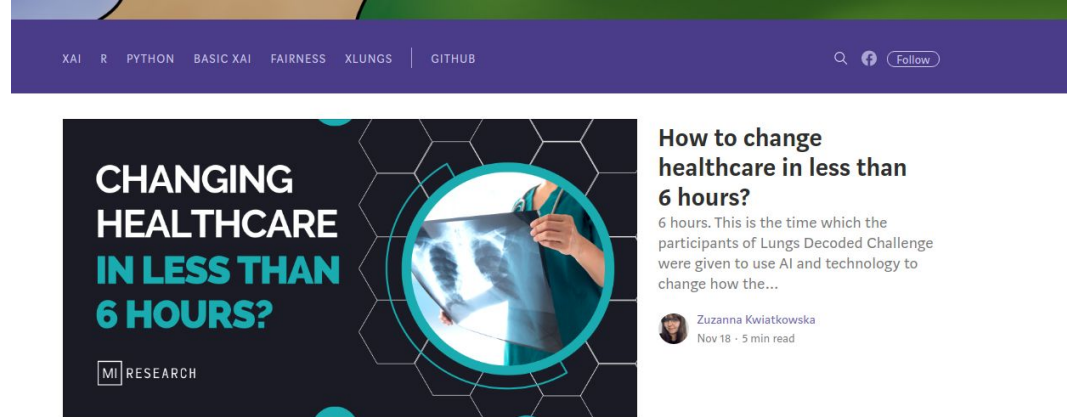
- [Welcome book, in Polish](#)
- [Welcome book, in English](#)
- Jak projekty są realizowane w labie
- [Ludzie w labie](#)
- Jak dostać maila w domenie @m12.ai
- Seminarium

Calendar

28	29	30	1 gru	2	3	4
○ 10:00 MI2 Seminar	○ 19:00 UNESCO sync	○ 10:00 Zajęcia z matematyki ○ 14:00 HOMER Seminar	○ 08:00 Enviromental policies weekly s ○ 10:00 Zajęcia z matematyki			

RML Blogs

- low frequency of appearing posts - once a month
- blog is beneficial to MI2
- blog may be to space to advertise your work and yourself



Division of duties in rooms

Now: cup system

To implement:

- reminder to water plants (slack)
- coffee and tea supply



Team building

- more informal meeting, more systematic



What can we do better to make things happen in the future?



Free proposals

Thank you for your attention!

TO DO:

- list of people with access to 44 -PB
- team beer -
- wiki information about Eden - HB
- Christmas meeting - 21st December, 14:00-..., room 44 - MK
- panel discussion about interdisciplinary - KW & PT

TO THINK ABOUT (1):

- peer coaching
- find people who write good blog and see important new things to share with audience (not duty but opportunity)
- do we need communication facility?
- hackathon

TO THINK ABOUT (2):

- more focus on guests on seminar
- improve the sound during seminar for remote attendees
- do we need more meetings like that?
- hackathon

REMEMBER:

- [calendar] add people to event who you expect to join
- [roadmap] if somebody is interested in research in lab, he can see roadmap