My first scientific paper in Machine Learning

Vadim Strijov

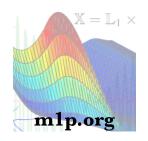
Moscow Institute of Physics and Technology

2021

The course produces student research papers

Machine learning projects: how to

- state the problem,
- make the project feasible,
- present results of the experiment



Science requires community:

- Student is a project driver, who wants to plunge into scientific research activities.
- Consultant, a graduated student, conducts the research and helps the student.
- Expert, a professor, states the problem and enlightens the road to the goal.

Lectures and seminars

Lecture

- ► Theoretical part: Machine learning for researchers
- ▶ Practice: The routine of scientific research and homeworks

Seminar

- Collective games
- Analysis of projects
- ▶ Talks

We read aloud every paper to prepare it for publication

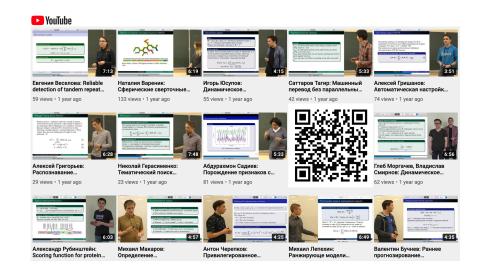
Four talks to convey your message to the audience

Week 3 Introductory pitch

- 6 The message
- 9 Computational experiment
- 12 Conference talk



YouTube channel: Machine Learning



Roadmap

- 1. Set the toolbox
- 2. Select your project
- 3. Read papers
- 4. Write introduction
- 5. State the problem
- 6. Set your experiment
- 7. Develop your theory
- 8. Make error analysis
- 9. Paper draft
- 10. Share your results
- 11. Finalize your paper
- 12. Present your talk

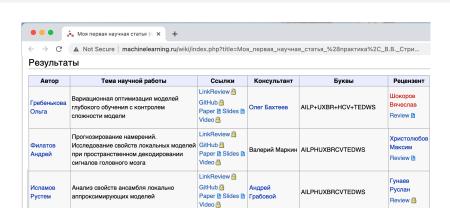


Deliveries are scheduled

- ► LinkReview with references and the literature review
- ► GitHub with the code and computational experiment
- ► Paper is ready for submission
- ► Slides for the presentation
- Video of the conference talk



Table with links to the deliveries and the marks



LinkReview A

Григорий

Марк Потанин

Paper P Slides Maлиновский

GitHub A

Video
LinkReview

GitHub A

Video A

Paper B Slides B

Ранее прогнозирование достаточного

Аддитивная регуляризация и ее

сетей глубокого обучения

метапараметры при выборе структуры

объема выборки для обобщенно линейной

Жолобов

Владимир

Вайсер

Кирилл

модели.

Вайсер

Кирилл

Review A

Жолобов

Владимир

Review A

AILPHUXBRCVTEWSF

All P+HUX+BBCV+TEDWS

New in 2021

News:

- Every your project is welcome!
- Your experts and consultants are highly appreciated!
- ▶ The number of students is not limited.

Challenges:

- Every student should work with the society!
- ▶ The project should be welcome to disseminate.
- The peer-review must be continuous.

A student paper example

Topic Models Selection for Reading Orders Generation over Document Collections Manusco K. E., Farenton E. F., Erenco M. A. nancer kelphyracit. rs., viareletrategs. rs., nanthirityander. rs. Manusch betished Physics and Technology. This paper investigates as impact of type makes in the quality of solding orders. A solding order superior superior for the quality of solding orders, A solding order angainst a softward of decreases in a tere have general to some openite decreases. A solding order or sold order Keyworks makes only probabilistic topic making making separa-The Market and the State of the

Type Melick Schooling for Bracking Orders Characters aren Dominant Collections

All the sizes in the control of the 2. Produces distanced:
A tending cold RPLAT over a document set. P is a directed acceler graph whose index energonal to the input decreases. In edges correspond to specificity relations among the chosensets. In particulars, a mode v, et V maps a successpect set D_c \(\tilde{\text{c}}\) of equivalent-obscurates. As slips v_c = v_c, between nodes v_c and v_c a figation flat, discussions behavior, the the energo-produced parameter set. D_c reverse decreases in behavior, to the energy-order set x_c. properties: (a) For each under $s_i \in T$ with D_i bring its corresponding set of documents, it holds that a climate of $d \in D$ major to s_i so $d \sim d$. In all $d \in D_i$.

Status S.S. France S. I., Srace S. I.

unishme cost of a regions of rhemotory consistent that convols one few into the older. An advantant to suppose and two electrics to be adjusted by S.

The other distances morely was consoled for Booding Order portions. We see not interested in low littless first term was see. We now for the relative concluding of such pair of discussments. To evaluate the time different based on the pairwise document solvings, we find fulfill the adjacency morely of for a tree operators using the following formula: To account the difference of two two represented by mutators A and \hat{A} , we use the norm reposed error (1995), which is defined as: 3 Model Relationships Between Documents 3.1 Topic Modeling control assumption of basic modeling in that the probability of the word or secure in $p(w\mid d) = \sum p(w\mid t)p(t\mid d) = \sum p_{to}R_{tb}$ The primary model (if proposed in [12] is the Peulahdibitic Latent Remarkie Analysis (PLAX). It obtains the substant matrices by maximizing the Bellinoid. $f_i(\theta,\theta) = \sum_{\theta \in \mathcal{H}} \sum_{n \in \theta} n_{n\theta} \log p(n \mid \theta) \rightarrow \max_{\theta \in \mathcal{H}}$ $\sum_{i \in \mathcal{G}} \varphi_{i+1} = \lambda_i \, \varphi_{i+2} \geqslant 0.$

Topi Multi-Science for Reading States Generalise and December Collections

 $\mathcal{L}(0,0) + \sum \mathcal{L}(0,0) \to \mathrm{span}.$ Spanity #. In order to make the µ_m distributions more space, the spanity regularizer is used. Market the following force: $J(0) = \sum_{n \in \mathbb{N}} \alpha_n \log \theta_n.$ $R = -\frac{c}{2}\sum_{n,k}\sum_{n,k}\sum_{n,m}g_{nm}g_{nm}$ The regularizers are multised to acquire some precise and wholde models. Supervisations of the model's who are 0 and 0 quantity values. However, the independently of plays is the most exception surgices are supervised measures because the entire of the control of estimated models independently of the assessor. In addition, ACTES beings new methods in hundling the models in Modelsine are vertices paid of test beingine are worship by a mission. Such are sufficiently and their indicates one help under the method of the desired to such help under the method of the desired to the high test of the desired to the desired to the desired to the high test of the desired to the desir building a better model.

Makkenskal ARTM solves the following optimization problem with old constraints: $\sum J_{n,i}(k_{n,i},0) + \sum J(i,j_{n},k_{n,i},0) \rightarrow \exp i.$ there is study for different modelline can be detained from the bests or their mets dates.

Finally, the Historical AHTM (AAHTM) medium here or more distant AHTM models for such level of the Houseley, label with each state. To take send-level inplies to be topics of the previous level, the application, discussed in [7] is used. $A = \sum_{i \in I} \sum_{n \in I} n_{in} \log \sum_{i \in I} \mu_{in} \phi_{in}.$ where T is the number of topics as the previous level, S is the number of topics as the next. Next, W in the standard in number, it is non-time of probabilities of next-level topics in previous-level case. This probabilities are obtained by solving the following problem:



Audior little of manipulal stratus over a document colorine is the ourseless. The resole certicules generates is discussed in [16] and [55] Bewere, they do not provide mobile relation, sub-trained provides.

 $\frac{1}{|S|} = \sqrt{n^{-1} L_{p}^{-1} d_{p}^{-1} + \frac{1}{|S|^{2}}} \, . \qquad (38)$ The application reasons that each performance look in the reasons sum of the area-lored topic, while the W marks can be whether in frather applications. All explainters and mobilities centained as applicable to the MMM mobility applications and problems of the contract of th

 $g(x) = -\sum R_{x} \log(R_{x}).$

Simulation dividings. The energy reads have a significant problem that makes it sensible if the discussion in feeding to high reads of the discussion in the last of the discussion in feedings to high reads affiliated problem from the last of the problem. We note that the discussion of the last of the las

began a six the first error began C. I have obbase the market Mr.

 $g_{ij}(\theta) = -\sum_{i}\theta_{ij}^{i}\sum_{j}\phi_{ij}\theta_{ij}^{j}\log\theta_{ij}^{j}.$ rel' rel'. Such water is more studie in femme of above assumptioned community. The effect in conjunced in the

5.3 Octaios: Distance is the second what parameter in reading order construction. It defines what documents result for read at the name time independently and what documents are not reput and should be read one after the other.

4 Computational Experiment

For our Dissistar Wildpedia documents and historicity of comparise to fit and validation our models. Both the Wildpedia notes and comparise descriptions are secondale through the Wilsoniak Research. Research, the comparise historicity in Wilsoniak is notified the term are the set of term. It is a commented anythin graph, assuing that from any uningary we can reach any is manifest, term liking our newls.
To validate the model we see this low with the Adulhoustines as root category and depth,
6, cantaining 1989 documents. On this set, all topic models were fixed and tone validated.
Additionally, to compare our newles with the initial [12], and explore the resideility of readpreparaments, we gathered the new with the rest of officialists. Incasings and depth 3.

Threat depending was performed we binamized for deciments, all bination was replaced with special of ODSELIA's taken.

Measure, to try the minimodal AHEM, we corrected biguous from the torce via the Buddies absorblish IM.

ANTH WILL SEPT MONEY 6.56

Algorithm I Lie Gries

p promoting Circles (A) do nigir matrix P, gravadity difference threshold is,

proceeding Circles (A) or a point matrix P, gravadity difference threshold is,

to be a construction of the construction of

Additive Engularization of Topic Models (AETM) availabled in [D], [B] and [B]bought used directly into topic modeling techniques. Still aiming to solve $\{4, AETM\}$ are explosions to manage the topic model quality, and the optimization periods has the form:

Names & S., Sanston K.V., Stewart M.A.

said it is imaged to a new tode x_i . Common bode x_i to the expansion node x_i .

6. Smarre bons, Ω off derivations belonging to R7. $C \mapsto \{x_i \in V_i | x_i \in A_i\}$ distributions belonging to R8. Sincia C has denotes D_i a.c.

8. Sincia C has denotes D_i a.c.

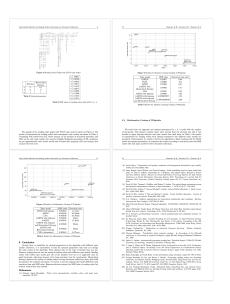
8. Sincia C has denotes D_i a.c.

8. Sincia C in bolds $x(i, d_i) \in M_i$, $d_i \in B_i$,

8. Sincia C in C in

5 Results of Experiments 5. Facility of Experiments
13. Machine Learning Carlining of Willipedia
Due to the sould not of the fations, it would be used for geometrie relevants. In Table 2
Electron 1. The control of the fations, it would be not for the fations of the fations of

A student paper example



Your benefits

- Research experience
- Published paper
- Qualification proof

Get info and ask your questions





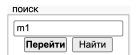


Russian: Вадим Викторович Стрижов, МФТИ

Автоматизация научных исследований

mlalgorithms@gmail.com

http://www.machinelearning.ru/wiki/index.php?title=m1



The project topics

Any topic is welcome on the open-access basis!