DRAFT
-
Explainable Federated Learning from Big Data over Networks: Fundamental Limits, Algorithms and Applications

PI: Alexander Jung

September 15, 2020

***Abstract*—Federated learning techniques have gained significant interest recently. These methods allow to train models from distributed local datasets. Federated learning tools are particularly attractive in medical applications as they do not require to share raw local data that might be confidential. Most existing methods for federated learning do not consider the intrinsic network structure of the local datasets. This project develops theoretical and practical tools for federated learning from big data over networks. To this end we will develop a theory of networked exponential families which have been proposed recently by the PI. Networked exponential families allow to jointly leverage on the information contained in local datasets and their network structure. These methods will be intrinsically privacy-preserving as they do not require to share local data, which might be sensitive. These methods exchange only model parameter updates between close-by nodes. We will also apply our recent information-theoretic approach to obtain explainable predictions. We illustrate the usefulness of our theory and methods by applying them to the high-precision management of pandemics. This project develops prospects to turn smartphones into personalized traffic-lights to optimally guide individual behaviour during pandemics. Personalized pandemic traffic-lights help humans to decide when to self-isolate or when to avoid certain places. The traffic-lights are controlled by localized predictive models, which are tuned by combining physical measurements (via sensing devices) and expert knowledge about diseases (via public health institutes).**

I. Aim and objectives

A. Significance of the research project in relation to current knowledge and the research-based starting points

How the project and the methods used are linked to previous international and/or national research (state of the art); Research premise, aims and objectives

The current coronavirus (COVID-19) pandemic has substantially affected our every-day life [6], [35]. There is an immediate threat to our health in the form of virus infections. What is even more, we are severely affected by non-pharmaceutical interventions (NPIs) [11]. The management of pandemics, including the timing of NPIs and optimal resource allocation for medicine and vaccines, relies on data.

AJ is with the Department of Computer Science, Aalto University, Finland.

We can model the data generated during a pandemic as a massive collection of local datasets that are related via network structures ("big data over networks") [8]. Indeed, local datasets are generated for individual humans by their smartphones and wearables [45]. These local datasets are related via several network structures. Contact networks relate datasets of humans that have been physically close to each other. Social networks relate datasets of humans with social ties [37]. Co-morbidity networks relate the local datasets of humans suffering from similar diseases [3].

One machine learning paradigm that aims at collections of local datasets is federated learning [28], [34]. Federated learning methods train machine learning models in a collaborative fashion without exchanging raw local data (see Figure 1)). As no local data is revealed, federated learning methods are attractive for sensitive applications, such as healthcare, which requires high levels of privacy protection [1].

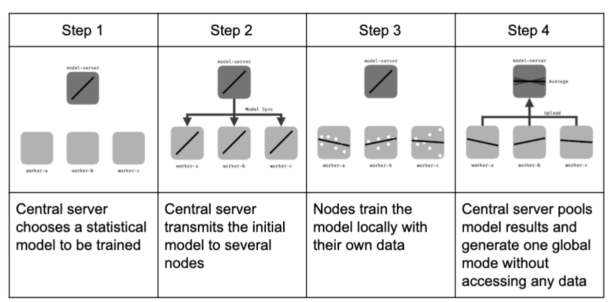


Fig. 1. Federated learning methods train a model from distributed local datasets. Attribution: Jeromemetronome / CC BY-SA (https://creativecommons.org/licenses/by-sa/4.0)

What sets this project apart from existing work on federated learning is the use of networked exponential families (nExpFam) as a novel modelling paradigm [19]. This paradigm uses a well-defined network structure inherent to the distributed local datasets. This project is driven by the hypothesis that big data over networks can be efficiently modelled using (nExpFam). Similar to probabilistic graphical models (PGM), nExpFam combine concepts from graph theory and probability theory to obtain tractable probabilistic models for massive datasets [29]. This project will use nExpFam as the main tool to obtain fundamental limits and efficient algorithms for federated learning in big data over networks.

A key obstacle for the success of federated learning is the heterogeneity of local datasets. Distributed local datasets often do not conform with an i.i.d. assumption [44]. Networked exponential families jointly model statistical and network structure of local datasets. This model allows us to study and develop federated learning algorithms that optimally leverage heterogeneous statistical properties of local datasets and their intrinsic network structure. Indeed,

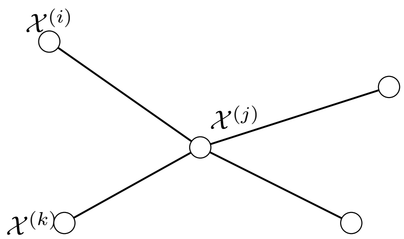
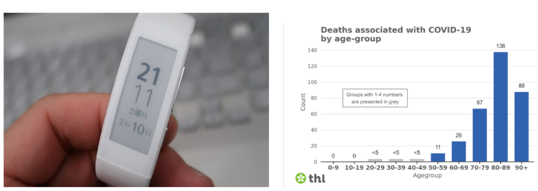


Fig. 2. A networked exponential family is a probabilistic model for local datasets $\mathcal{X}^{(i)}$ that are organized in a network.

This project models local datasets as realizations of high-dimensional random variables. The probability distribution of these random variables are assumed to belong to an exponential family [9], [30], [42]. Exponential families are parametrized sets of probability distributions with attractive statistical and computational properties [50]. Special cases of exponential families are the multivariate normal distribution and the multinomial distribution [7], [50]. These special cases can be used to model the measurements of wearables or histograms of random samples (see Figure 3).



techno_m5 / CC BY-SA
(https://creativecommons.org/licenses/by-sa/2.0)

https://thl.fi/fi/

Fig. 3. (Left) The statistical properties of wearable measurements can be modelled using a multivariate normal distribution. (Right) The statistical properties of histograms computed for sub-populations can be modelled using a multinomial distribution.

We assume local datasets are generated from distributions belonging to the same exponential family. However, we allow the parameters of the distributions to vary between individuals. An nExpFam couples the distributions of the local datasets by requiring the parameters of well-connected individuals to be similar. The methods obtained from nExpFam therefore allow to jointly leverage on the information contained in network structure between and statistical properties of local datasets. Loosely speaking, we can lend and borrow statistical strength of the local datasets between neighbours.

The main subject of this project are the statistical and computational aspects of nExpFam as a probabilistic model for networked data. The statistical aspects include conditions on the network structure and local datasets such that the underlying nExpFam can be accurately learnt. Similar in sprit to complex networks [4], [37], we will study the interplay between local network structure and emergence of global properties of the nExpFam. We will illustrate the usefulness of our theory of nExpFam by applying the resulting methods to the high-precision management of pandemics. Our methods will be used to optimally steer NPIs and allocate resources such as drugs and vaccines.

The computational aspects of nExpFam include the design of efficient methods to learn the parameters of an nExpFam. To this end, we will apply distributed convex optimization methods for nExpFam. These methods cope with limited computational resources and do not share sensitive local data but only parameter estimates. In contrast to existing federated learning methods [44], we do not assume server-client architecture but a fully decentralized implementation.

The computational infrastructure required by our methods is similar to those underlying contact tracing apps [10]. Tracing apps use near-field communication between smartphones to detect potentially risky contacts. In contrast, our methods use near-field communication between smartphones to update the local parameters of personalized predictive models. These updates only requires exchange of parameter values but no raw data that might be sensitive is revealed to others.

The acceptance of machine learning methods often depends crucially on their explainability, interpretability and transparency [36], [41], [43]. Explainable (or interpretable) methods provide human users with means to understand (or comprehend) why a particular prediction has been made. Loosely speaking, explainable methods provide answers to the question "What are the reasons for obtained this particular prediction?" [43]. We will extend our recent line of work [20], [24] on explainable machine learning to federated learning from big data over networks.

Our approach to explainability models the user background using some feed signal which we refer to as user summary. This user summary reflects the intuition of the ML user for how the predictions should behave. This intuition might be shaped by the media and the information obtained from policy makers.

We will verify our theoretical findings using their application to high-precision management of pandemics (see Figure 5). The current coronavirus (COVID-19) pandemic has substantially affected our every-day life [6], [35]. There an immediate threat to our health in the form of virus infections. What is even more, we are severely affected by the related non-pharmaceutical interventions (NPIs) [11]. Until the availability of a safe and effective vaccination and medication, NPIs might be the only effective tool to manage pandemics and avoid collapse of our healthcare infrastructure [38].

Since NPIs often come with massive collateral socio-economic damages they must be used carefully [11], [39].

During the onset of the Covid-19 pandemic, NPIs were typically enforced for entire countries. Clearly, such a course-grained approach ignores the significant regional variability in the epidemic situation.

To minimize collateral damage, most countries have adjusted their policies and aim at applying NPIs more locally such as closing only individual schools. The Austrian government has introduced a coronavirus “traffic light” system which evaluates the situation individually for different districts (see https://corona-ampel.gv.at/karte/).

This project takes the idea of a coronavirus traffic light one step further to the finest level of granularity, i.e., individual humans. We will develop theory and methods for a personalized coronavirus traffic-light system. This system provides tailored recommendations for individuals if to self-isolate or avoid certain places (e.g., elderly people homes).

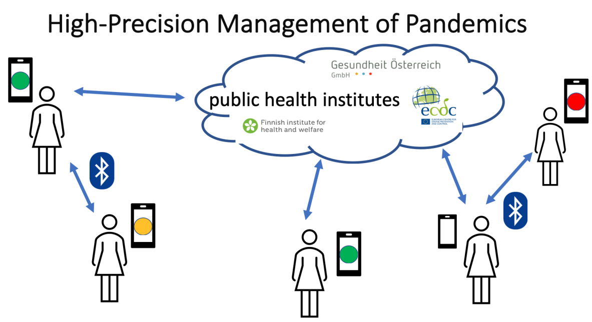


Fig. 4. A peronsalized coronavirus traffic-light system.

The enforcement of NPIs is based on the predictions (or forecasts) obtained from epidemiological models [12], [15] (see also https://github.com/cdcepi/COVID-19-Forecasts). Existing models for disease spread are calibrated using statistics about larger populations, e.g., on the level of municipalities or entire countries [47]. In contrast, this project develop methods that use real-time data about individual humans observed via their smartphones and other wearables [13], [45].

This project develops methods that combine, in a privacy-preserving way, the raw data collected on smartphones or wearables with data and knowledge provided by public health institutes (see Figure 5). We design ethical recommendation systems that also take into account the socio-economic consequences of its recommendations [16]. The recommendations for an individual might be different depending if the person is working in a system-critical role (teacher, nurse or heatlh-care policy maker).

B. Research questions and/or hypotheses

This project is driven by a main working hypothesis that nExpFam are a suitable modelling paradigm for federated learning from massive networked-structured data (big data over networks). In particular,

- We *can efficiently model* big data over networks arising in important application domains, such as management of pandemics, using networked families (nExpFam).
- Fundamental limits of distributed federated learning from big data over networks *are reflected* in the intrinsic properties of nExpFam.
- Distributed optimization method for learning nExpFam *can be used* to design distributed federated learning algorithms for computing personalized predictions.
- Our recently proposed information-theoretic approach to explainable machine learning *allows to* design transparent and explainable federated learning methods.

C. Expected research results

and their anticipated scientific impact, potential for scientific breakthroughs and for promoting scientific renewal; Research impact within the scientific community; Project’s novelty or added value for science

This project will significantly extend the state of the art in federated learning theory and methods. This project will develop a novel modelling paradigm for studying statistical and computational aspects of explainable federated learning from big data over networks. This novel modelling paradigm will be based on our recent work on networked exponential families (nExpFam) and an information-theoretic approach to explainable machine learning.

We expect three main results of this project.

- A characterization of the statistical properties of federated learning methods. These properties include fundamental limits and trade-offs between computational resources, data, privacy and explainability (or interpretability). These fundamental limits will be used to certify the (sub-)optimality of distributed optimization methods for learning nExpFam.
- Design of efficient algorithms for federated learning from big data over networks.
- Construction of explanations for the local predictions provided by federated learning methods.

These results will be applied to the high-precision management of pandemics by predicting optimal NPIs on the finest level of granularity, i.e., individuals. Therefore, this project also contributes to the development of scientific disciplines related to epidemics and health-care systems.

II. Implementation

A. Work plan and schedule

Detailed description of the research to be performed, starting from objectives, scientific references and preliminary data (if available); Description of research tasks, their implementation and interconnections; If necessary, description of the responsibilities and management related to these tasks; Schedule for project implementation, incl. research tasks and work packages, distribution of personnel resources, and project milestones and deliverables

We organize the research required to achieve the expected project result as four inter-dependent research tasks (RT). Each of the four RT roughly covers one of the four hypotheses underlying this project.

RT1 Networked Exponential Families for Federated Learning. We model local datasets arising in applications such as the management of pandemics as realizations of random

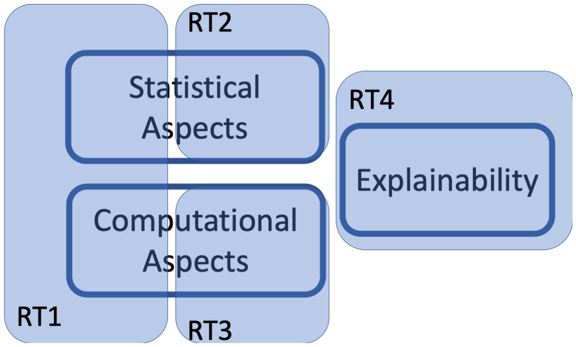


Fig. 5. The project research is organized in four interrelated research tasks.

processes. The local dataset $\mathcal{X}^{(i)}$ of some individual i is distributed according to an exponential family

$$p(\mathcal{X}^{(i)}; \mathbf{w}^{(i)}) := b^{(i)}(\mathcal{X}^{(i)}) \exp\big((\mathbf{w}^{(i)})^{T} \mathbf{t}^{(i)}(\mathcal{X}^{(i)}) - \Phi^{(i)}(\mathbf{w}^{(i)})\big). \quad (1)$$

The exponential family (1) is parametrized set of probability distributions. Probability distributions of the form (1) are appealing statistically and computationally [50]. These distributions are distinct in the sense of maximizing the entropy given knowledge of some ("global") statistics which we stack into the vector $\mathbf{t}^{(i)}(\mathcal{X}^{(i)})$.

The entries of $\mathbf{t}^{(i)}(\mathcal{X}^{(i)})$ might be empirical averages for local datasets $\mathcal{X}^{(i)}$ obtained from time series (from wearables). There is a large body of work on computationally efficient methods for learning (fitting) the parameters $\mathbf{w}^{(i)}$ to the observed data $\mathcal{X}^{(i)}$ [50].

Note that we allow for different weight vectors $\mathbf{w}^{(i)}$ in (1) for each individual $i \in \mathcal{V}$. What is more, and clearly distinguishes nExpFam from PGM [29], [50], we even use separate probability spaces to model the local datasets $\mathcal{X}^{(i)}$.

It is a main hypothesis of this project that it is not useful to use the same probability space for all observed data. Using this approach (which is the main point of PGM) would require us to model statistical dependencies between local datasets on a microscopic level which we consider impractical.

We couple the distributions of the local datasets $\mathcal{X}^{(i)}$ via an empirical graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$. Nodes $i \in \mathcal{V}$ represent individuals that are connected by links or edges. Each edge $\{i, j\}$ is assigned a positive weight $W_{i,j}$ representing the level of similarity between individuals $i, j \in \mathcal{V}$.

The network structure (edge set) of the empirical graph $\mathcal{G}$ is obtained by different means. Contact networks arise from physical proximity and they are effectively known as each individual knows the close-by individuals via the near-field communication ("bluetooth") of the their smartphones. Social networks are available from phone address books and public registers. We can also obtain networks by connecting people via biological similarities such as the same blood type [46].

This task launches from our recent work on special cases of nExpFam obtained for simple network structures and linear models [2], [22], [25], [49]. We will extend this work to more complicated network structures and larger classes of exponential families including time series models.

A focus will be on nExpFam defined on multi-layer networks that might be used to jointly represent contact, social and genetic networks. The results of this task will serve the long-term goal of developing an information-theory of nExpFam. This theory combines the information geometry of conventional exponential families with the discrete geometry of the empirical graph $\mathcal{G}$.

We will verify the usefulness of nExpFam by applying them to the high-precision management of pandemics. Most existing network approaches to study pandemics use low-dimensional models for the local data of individuals [37]. These models often reduce the local dataset to a single categorical variable, taking on values "susceptible", "infected" and "immune" [47]. In contrast, nExpFam models local datasets as high-dimensional stochastic processes which are distributed according to an exponential family (1).

In contrast to existing low-dimensional models, nExpFam allow for more fine-grained modelling for the temporal dynamics of symptoms or changes in the behaviour of individuals. The temporal dimension might be instrumental to distinguish between a harmless cold and a more serious infection [33].

RT2 Fundamental Limits and Tradeoffs.

This task aims at using the intrinsic properties of nExpFam (see RT1) to characterize fundamental limits for accurate learning of nExpFam. In order to learn nExpFam we need use the information provided by observing the local datasets and the network structure of the empirical graph. In general we can only observe the local datasets for a subset of individuals only. We refer to this subset as the sampling set.

How much local data do we need ? How much do we need to know about the structure of the empirical graph ?

These theoretical results will be applied to resource allocation (testing, computation time for getting predictions, vaccination) in the management of pandemics.

Our focus is on conditions on the sampling set and network structure such that accurate predictions can be learnt given resource constraints. These conditions will allow to optimally allocate resources such as testing capacities or vaccinations. We can also use these theoretical insights to decide which group of people should be prioritized for vaccination campaigns. This task will take off from our recent work on conditions ensuring that networked linear models can be learnt accurately [21], [23], [26]. We will extend our previous work to larger subclasses of exponential families, such as Gaussian Markov random fields modelling the dependencies between different parameters.

Another main goal of this task is the derivation of conditions that allow to learn network structure in a data driven fashion. We have recently considered the problem of learning network structure of probabilistic graphical models (PGM) [18], [48]. Both, PGM and nExpFam are probabilistic models that use graph-theoretic concepts for managing complexity.

In contrast to PGM, within nExpFam we do not model statistical dependencies or correlations between the local datasets as this is typically intractable. This task will explore the relations between PGM and nExpFam that would allow to make use of PGM methods to learn nExpFam. As another launching point, we will use concepts from network flow optimization to learning the structure of the empirical graph of a networked dataset [21].

RT3 Distributed Federated Learning Algorithms. Learning nExpFam from networked data can be formulated as (convex) optimization problems [23], [25], [26], [49]. The formulation as optimization problems provides us a vast algorithmic toolbox for federated learning in the form of distributed convex optimization methods [5], [40].

While RT2 is focused on the statistical aspects of nExpFam, this task studied computational aspects of learning nExpFam. In general, the statistical and computational aspects of machine learning models are closely related. As a point in case, we note that fixed-point characterizations of the solutions to learning problems lends naturally to iterative algorithms via fixed-point iterations.

In this task we develop novel algorithms for distributed federated learning of models from distributed local datasets [44]. We propose a novel design paradigm for federated learning algorithm which is rooted in convex optimization for nExpFam models. The resulting algorithms are robust to noisy data and changing network topologies such as arising from link failures. This task will take off from our recent work on primal-dual optimization methods for distributed learning of networked linear models [21], [23], [26]. These message passing methods are an instance of the federated learning paradigm .

RT4 Explainable Federated Learning. The explainability (or transparency) of machine learning methods increasingly becomes a legal requirement [14]. Moreover, the acceptance of ML methods by humans seems to depend crucially on their explainability [27], [31]. We will extend our recent information-theoretic approach to explainable ML to construct tailored explanations of predictions [20], [24].

By combining our recent information-theoretic approach [20], [24] to explainable ML with the results of RT1 and RT3, we will develop explainable federated learning methods. These explainable learning methods take into account the heterogeneous background of different users. We will develop different methods for constructing explanations or regularizing the learning process such that the resulting predictions are maximally comprehensible.

B. Research data and material, methods, and research environment:

Research data to be used, justifications and information on data collection or acquisition, data analyses and use of data, taking into account issues such as intellectual property rights; Research methods and how they will contribute to answering the research questions or confirming the hypotheses, or how they will support the chosen approach; Description of local, national and/or international research environment including research infrastructures. Enter the infrastructures to be used also on the tab 'Affiliations' in the online services.

Research Data. The project is mainly theoretic in nature. In particular RT1, RT2 and RT3 can be completed with synthetic data only. For RT4 we will conduct small-scale experiments involving some group members (with their informed consent) as well as volunteers that will be recruited from the massive courses taught by the PI.

A main component of the project the application of our theoretic results in the management of pandemics. To this end, we will use open epidemic data and statistics as well as open data on demographics such as geographical population distribution and density. This data is provided by various public institutes such as the European Center for Disease Control (https://www.ecdc.europa.eu/en/) or national health-care institutes (https://thl.fi/fi/). We will also use data provided by statistical offices such as Statistics Finland (https://www.stat.fi), Statistics Austria (http://www.statistik.at/) or Eurostat (https://ec.europa.eu/eurostat/de/home). This open data will be used to construct realistic models for the network structure of individuals. To model the symptoms caused by infections diseases (see RT1), we will mainly rely on scientific publications [32], [33].

Research Methods. The project will be implemented using a mix of theoretical analysis and experimental validation in numerical experiments. The main tools for the theoretical analysis will come from convex optimization, distributed non-linear systems, probability theory and information theory. The numerical experiments will be implemented mainly in Matlab and Python and according to the principle of reproducible research. To study the explainability of the develop techniques we will also run studies on small groups of volunteers. These volunteers will be recruited from different domains such as research groups and students of the massive courses taught by the PI.

Research Infrastructure. The main research environment is the Department of Computer Science at the Aalto University School of Science. With over 40 professors and more than 400 employees, the department is the largest Computer Science unit in Finland and facilitates fruitful interaction between different aspects of computer science. The department is ranked among the top 10 Computer Science departments in Europe and in the top 100 globally. ICT and digitalization is also one of the seven key research areas of Aalto University. The department leads the Academy of Finland Flagship Finnish Center for Artificial Intelligence (FCAI, https://fcai.fi/) and has a joint research institute with University of Helsinki, the Helsinki Institute for Information Technology (HIIT, https://www.hiit.fi/).

In addition to standard resources provided by the department, there is support for midrange HPC and data management through the Science-IT project of the Aalto University School of Science. Science-IT and its flagship HPC unit Triton are a part of the national Finnish Grid and Cloud Infrastructure (FGCI). Currently, Triton provides 9300 computing cores and over 30 servers for GPU computing. The CS department IT offers hands-on support and training, and software research engineers available for research projects. Researchers have also free access to national CSC resources. These include e.g. upcoming Europe's most powerful GPU based supercomputer LUMI.

The courses on machine learning and artificial intelligence taught by the PI have been enrolled by several thousands of students from all over (and beyond) Aalto. The results of this project will be disseminated partly via these courses and related student projects. The visibility of the PI via his teaching will also be instrumental for recruiting talented students for spin-off thesis projects.

III. Applicant, possible research team and collaborators

A. Project personnel and their project-relevant merits:

Tasks, roles and key merits of the project PI and the project's researchers Names and/or level of education of the project's researchers (if known) How the project is linked to previous or other research by the applicant How the project advances the research career of the applicant

The PI Alexander Jung is currently an Assistant Professor (level 2) for Machine Learning at the Department of Computer Science at Aalto University. His research revolves around fundamental limits and efficient methods for massive datasets. Since recently, his focus is on combining tools from machine learning and complex network concepts to efficiently process network structured data. Together with his collaborators, he pioneered the characterization of network structure and statistical properties of data that allow for accurate learning from big data over networks. The excellence of his research work is documented by numerous publications in top-tier journals. He is the first author of a paper that received a Best Student Paper Award at the premium signal processing conference IEEE ICASSP in 2011. In 2018, he was awarded an Amazon Web Services (AWS) Machine Learning Award. Prof. Jung currently serves as an Associate Editor for the IEEE Signal Processing Letters. He has been chosen as Teacher of the Year by the Department of Computer Science in 2018.

B. Collaborators and their project-relevant key merits:

National and international collaborators of key significance to project implementation as well as their merits Justifications for the collaborators, description of what is achieved through the collaboration

IV. Responsible science

A. Research ethics:

Information on ethical issues (e.g. ethical governance procedures, informed consent, anonymity of subjects and withdrawal from research) that concern the chosen topic, methods and data; Information on research permits granted or pending; Read more in the ethical guidelines.

It is noteworthy that no studies will be conducted that may raise ethical issues, such as research carried out with human embryos or other human cell types, human beings or animals, genetic information, or personal data. In the unlikely case that ethical issues arise, Aalto University Research Ethics Committee will be consulted at need. We would also like to note that Aalto University (and thus also the project researchers) is committed to following the guidelines issued by the Finnish Advisory Board on Research Integrity on good scientific practice, handling violations against it, and valid legislation. As in all externally funded research projects of Aalto University, the source code will be owned by the university but the aim is to publish the code as openly as possible after possible exploitation activities (e.g., patenting) have been thoroughly analyzed. If public datasets are utilized, the corresponding sources will be cited as part of good scientific practice.

B. Equality and non-discrimination

Information on how the project will promote equality and non-discrimination within itself or in society at large.; Read more in the equality and non-discrimination guidelines.

We are committed to the Aalto Code of Conduct in the university environment that declares that members of Aalto University will behave responsibly and respect the rights of others. Each member of the community has a right to be treated with respect regardless of their gender, age, ethnic or national origin, nationality, religion, or other personal characteristics. All the research group members will be selected solely based on scientific qualifications and competence. The gender balance in a research group will be maintained with due emphasis. The PI takes the promotion of equality and non-discrimination within academia very seriously. In particular, he always tries to fill research or teaching assistant positions always balanced with women and men (given a sufficient number of suitable candidates). For the Aalto spearhead course Machine Learning: Basic Principles he managed to hire 8 females out of a total of 21 TAs during the course edition 2018 (with more than 800 students enrolled).

C. Open science:

Publication plan that supports open access (Academy-funded projects are required to commit to open access publishing) • Read more in the open science guidelines; Brief plan for data management: how the data will be stored during the project, how any legal and ethical issues related to data distribution will be resolved, and where the data will be made available after the end of the project. Funding recipients must submit an actual data management plan within eight weeks of the funding decision. The payment of the funding is conditional on the submission of the plan. ; Read the guidelines on the data management plan.

The outputs of research will be mainly scientific articles containing analytic results and their derivation (proofs). Some of the theoretical findings will lead to efficient algorithms which will be implemented in big data frameworks. The source code of these implementations will be made publicly available to the extent possible in order to facilitate reproducible research.

All publications published within this project will be made open access on minimum SHERPA/RoMEO green level through Aalto ACRIS Research Information Portal for pre- and post-print manuscripts https://research.aalto.fi/en/. Funding is also applied to cover Open Access publication fees. Research data will be made available openly. At all stages, Aalto University Open Access Policy will be followed (see http://libguides.aalto.fi/openaccess).

The source code will be written in various scientific programming languages (e.g., MATLAB, C++, Python). We estimate that approximately 10000 lines of code will be written and published in Git during the entire project. The source code will be extensively commented to ensure that the code is as intelligible as possible and can be utilized by others after it is published. We will handle versioning of the code with the help of Git's inbuilt features (e.g., the branching model).

Data sharing and long-term preservation. The scientific articles produced in the project will be stored in ArXiv https://arxiv.org/ and AaltoDoc https://aaltodoc.aalto.fi. The developed source code of the algorithms will be stored in Git https://git-scm.com/. Aalto University file services will be used during the project. The services include a snapshot feature and regular backups that generate file versions automatically to recover fromunwanted deletions - tape backups provide also system-level disaster recovery. All the Aalto University laptops utilized in the project include automatic data encryption with Bitlocker, and secure file transfer over the network with a VPN solution.

The source code will be openly shared via GitHub (with a free license). Via GitHub, anyone can access the code, modify it with standard tools (a computer and a C++compiler) and use it as they wish. The source code will be linked to scientific publications through which the potential users will find out about it. We will also actively promote the source code in other fora, such as Bitbucket.

D. Risk assessment and alternative implementation strategies

Critical points for success, probability of risks, means by which risks can be managed, and alternative implementation strategies

The main conceptual risk stems from the validity of our main hypothesis that networked data can be efficiently modelled using nExpFam. It might turn out ignoring the statistical dependencies (correlations) between local datasets is not justified in most applications, including pandemics. We can manage this risk by adapting the methods using concepts from the theory of probabilistic graphical models.

A main operational risk is posed by the recruitment of suitable graduate and post-graduate students. If it is not possible to find a suitable candidate, we will re-organize the research tasks such that they can be carried out by several undergraduate projects.

E. Sustainable development objectives:

Brief description of how the project can promote one or more of the eight goals for sustainable development: equal prospects for wellbeing, a participatory society for citizens, sustainable employment, sustainable society and local communities, a carbon-neutral society, a resource-wise economy, lifestyles respectful of the carrying capacity of nature and decision-making respectful of nature; Read more in the sustainability guidelines.

The project results (theoretical results and developed algorithms) will promote sustainable development in several aspects.

Equal prospects for wellbeing. We develop methods that are accessible to anybody with a smartphone. This evens out differences in access to traditional channels of healthcare (human experts). Moreover, we expect our methods that support tailored NPI measures to ease the burden of collateral damages caused for underprivileged people (e.g. isolation in a 20 sqm apartment with three kids is more challenging than isolation in 200 sqm villa).

Lifestyles respectful of the carrying capacity of nature and decision-making respectful of nature. A positive side effect of personalized NPIs such as recommended self-isolation and therefore not using fossil fuel based means of transport, is a contribution towards a carbon-neutral society. Personalized NPIs will also help to optimally use face masks, which in turn reduces the waste generated by masks. The explanations provided for personalized NPIs will nudge individuals to become more aware of and respectful for the intrinsic capacity of nature (which also includes other humans!) [17].

Sustainable Employment. Having tailored NPI instead of country-wide lock downs also helps to avoid significant fluctuations on the job market e.g. for restaurants or hotels.

A participatory Society for Citizens. We develop explainable machine learning methods that provide tailed explanations for the individual predictions. These explanations are tailored to maximally enable users with different backgrounds and level of education. We believe that explainability will also motivate individuals to reflect about their behaviour and its effect to the well-being of others.

Sustainable society and local communities. A key risk during a pandemic is the overload of infrastructure, such as healthcare. The avoidance of such an overload is particularly challenging in small communities (e.g. small island with few health care personal). The methods developed in this project will make small communities more resilient during pandemics by minimizing the risk of infrastructure overload.

V. Societal effects and impact

A. Effects and impact beyond academia:

Brief description of the appeal, utilisation potential and application areas of the research results beyond the scientific community; For instance, provide a self-assessment of the expected societal impact of the research in the long or short term. Impact beyond academia may come in many different forms depending on the research field and the project. For example, science is a source of wealth and prosperity, but it also improves our understanding of the world and enhances the level of civilisation, supports the development of good practices and informs decision-making; Read more about the broader impact of research.

The research results will contribute to allowing policy-makers use NPIs on the finest possible level of granularity, i.e., individual persons. It is therefore hard to underestimate the impact of our research project on the well-being (in the most general sense) of our societies.

[49]

VI. Conclusion

References

[1] N. Agarwal, A.T. Suresh, F. Yu, S. Kumar, and H.B. McMahan. cpsgd: Communication-efficient and differentially-private distributed sgd. In *Proc. Neural Inf. Proc. Syst. (NIPS)*, 2018.

[2] H. Ambos, N. Tran, and A. Jung. Classifying big data over networks via the logistic network lasso. In *Proc. 52nd Asilomar Conf. Signals, Systems, Computers*, Oct./Nov. 2018.

[3] A. Barabási, N. Gulbahce, and J. Loscalzo. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12(56), 2011.

[4] A.-L. Barabási and M. Pósfai. *Network science*. Cambridge Univ. Press, Cambridge, UK, 2016.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, volume 3. Now Publishers, Hanover, MA, 2010.

[6] Z.J. Cheng and J. Shan. 2019 novel coronavirus: where we are and what we know. *Infection*, 48(2):155–163, 2020.

[7] L. Clifton, D.A. Clifton, M.A.F. Pimentel, P.J. Watkinson, and L. Tarassenko. Gaussian processes for personalized e-health monitoring with wearable sensors. *IEEE Transactions on Biomedical Engineering*, 60(1), 2013.

[8] S. Cui, A. Hero, Z.-Q. Luo, and J.M.F. Moura, editors. *Big Data over Networks*. Cambridge Univ. Press, 2016.

[9] G. Darmois. Sur les lois de probabilites a estimation exhaustive. *C. R. Acad. Sci. Paris*, 1935.

[10] Eidg. Department des Inneren EDI. Faktenblatt - swisscovid app hilft das coronavirus in schach zu halten. Technical report, Bundesamt fuer Gesundheit BAG, Jun. 2020.

[11] S. Flaxman et.al. Estimating the number of infections and the impact of non-pharmaceutical interventions on covid-19 in 11 european countries. Technical report, Imperial College London, doi:https://doi.org/10.25561/77731., 2020.

[12] Neil M. Ferguson, Derek A. T. Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, 2006.

[13] K. Guk, G. Han, J. Lim, K. Jeong, T. Kang, E.-K. Lim, and J. Jung. Evolution of wearable devices with real-time disease monitoring for personalized healthcare. *Nanomaterials*, 9(6), Jun. 2019.

[14] P. Hacker, R. Krestel, S. Grundmann, and F. Naumann. Explainable AI under contract and tort law: legal incentives and technical challenges. *Artificial Intelligence and Law*, 2020.

[15] E. Hunter, B. Mac Namee, and J. Kelleher. An open-data-driven agent-based model to simulate infectious disease outbreaks. *PLOS ONE*, 13(12), 2018.

[16] A. Jobin, M. Ienca, and E. Vayena. The global landscape of ai ethics guidelines. *Nature Machine Intelligence*, 1(9):389–399, 2019.

[17] T.E. Jovejoy. To prevent pandemics, stop disrespecting nature. *National Geographic*, May 2020.

[18] A. Jung. Learning the conditional independence structure of stationary time series: A multitask learning approach. *IEEE Trans. Signal Processing*, 63(21), Nov. 2015.

[19] A. Jung. Learning networked exponential families with network lasso. *arXiv*, 2019.

[20] A. Jung. Explainable empiricial risk minimization. *submitted to IEEE Sig. Proc. Letters (preprint: https://arxiv.org/pdf/2009.01492.pdf)*, 2020.

[21] A. Jung. On the duality between network flows and network lasso. *IEEE Sig. Proc. Lett.*, 27:940 – 944, 2020.

[22] A. Jung, A. Heimowitz, and Y. C. Eldar. The network nullspace property for compressed sensing over networks. In *Sampta*, Tallinn, Estonia, Jul. 2017.

[23] A. Jung, A O. Hero, A. Mara, S. Jahromi, A. Heimowitz, and Y.C. Eldar. Semi-supervised learning in network-structured data via total variation minimization. *IEEE Trans. Signal Processing*, 67(24), Dec. 2019.

[24] A. Jung and P.H.J. Nardelli. An information-theoretic approach to personalized explainable machine learning. *IEEE Sig. Proc. Lett.*, 27:825–829, 2020.

[25] A. Jung, N.T. Quang, and A. Mara. When is Network Lasso Accurate? *Front. Appl. Math. Stat.*, 3, Jan. 2018.

[26] A. Jung and N. Tran. Localized linear regression in networked data. *IEEE Sig. Proc. Lett.*, 26(7), Jul. 2019.

[27] J. Kagan. Motives and development. *Journal of Personality and Social Psychology*, 22(1):51–66, 1972.

[28] P. Kairouz, B. Mcmahan, B. Avent, A. Bellet, and M. Bennis. Advances and open problems in federated learning. *HAL hal-02406503*, 2019.

[29] D. Koller, N., and Friedman. *Probabilistic Graphical Models: Principles and Techniques*. Adaptive computation and machine learning. MIT Press, 2009.

[30] B. Koopman. On distributions admitting a sufficient statistic. *Trans. Amer. Math. Soc.*, pages 399–409, 1936.

[31] A.W. Kruglanski and D.M. Webster. Motivated closing of the mind. *Psychol. Rev.*, 103(2), 1996.

[32] F.Y. Lan, R. Filler, and et.al. S. Mathew. Covid-19 symptoms predictive of healthcare workers' sars-cov-2 pcr results. *PLOS ONE*, 15(6), Jun. 2020.

[33] J.R. Larsen, M.R. Martin, J.D. Martin, P. Kuhn, and J.B. Hicks. Modeling the onset of symptoms of covid-19. *Frontiers in Public Health*, 8:473, 2020.

[34] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of Machine Learning Research*, volume 54, pages 1273–1282, Fort Lauderdale, FL, USA, Apr. 2017.

[35] S. Moisio. State power and the covid-19 pandemic: the case of finland. *Eurasian Geography and Economics*, pages 1–8, 2020.

[36] C. Molnar. *Interpretable Machine Learning - A Guide for Making Black Box Models Explainable*. [online] Available: https://christophm.github.io/interpretable-ml-book/., 2019.

[37] M. E. J. Newman. *Networks: An Introduction*. Oxford Univ. Press, 2010.

[38] G. Nedjati-Gilani et al. N.M. Ferguson, D. Laydon. Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. Technical report, Imperial College London, Mar. 2020.

[39] A. Pak, O.A. Adegboye, A.I. Adekunle, K.M. Rahman, E.S. McBryde, and D.P. Eisen. Economic consequences of the covid-19 outbreak: the need for epidemic preparedness. *Frontiers in Public Health*, 8, 2020.

[40] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):123–231, 2013.

[41] S. Lo Piano. Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanities and Social Sciences Communications*, 7(1):9, 2020.

[42] E. Pitman and J. Wishart. Sufficient statistics and intrinsic accuracy. *Mathematical Proceedings of the Cambridge Philosophical Society.*, 32(4):567–579, 1936.

[43] M.T. Ribeiro, S. Singh, and C. Guestrin. "Why should i trust you?": Explaining the predictions of any classifier. In *Proc. 22nd ACM SIGKDD*, pages 1135–1144, Aug. 2016.

[44] F. Sattler, K. Müller, and W. Samek. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE Transactions on Neural Networks and Learning Systems*, 2020.

[45] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl. Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. *IEEE Transactions on Mobile Computing*, 13(4):907–920, April 2014.

[46] J.C. Silva-Filho, C. Germoglio Farias de Melo, and J. Lima de Oliveirac. The influence of abo blood groups on covid-19 susceptibility and severity: A molecular hypothesis based on carbohydrate-carbohydrate interactions. *Med Hypotheses*, Nov 2020.

[47] S. Thurner, P. Klimek, and R. Hanel. A network-based explanation of why most covid-19 infection curves are linear. *Proceedings of the National Academy of Sciences*, 2020.

[48] N. Tran, O. Abramenko, and A. Jung. On the sample complexity of graphical model selection from non-stationary samples. *IEEE Transactions on Signal Processing*, 68:17–32, 2020.

[49] N. Tran, H. Ambos, and A. Jung. Classifying partially labeled networked data via logistic network lasso. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3832–3836, Barcelona, Spain, May 2020.

[50] M. J. Wainwright and M. I. Jordan. *Graphical Models, Exponential Families, and Variational Inference*, volume 1 of *Foundations and Trends in Machine Learning*. Now Publishers, Hanover, MA, 2008.