Università
della
Svizzera
italiana

Faculty of Informatics
Master's Research Scholarships (MaRS) Projects
2024

[MaRS-2024-01]
Scalable State Machine Replication
Contact: Prof. Fernando Pedone
State Machine Replication (SMR) is a well-established replication technique used by many production systems, including Apache Zookeeper, Google Chubby, Windows Azure storage, Google Spanner, and many others. Scalable State Machine Replication (S-SMR) is a recent extension of SMR developed at the distributed systems group at USI that promises unlimited performance in addition to configurable fault tolerance. Some initial efforts, for example, resulted in a prototype that outperforms Zookeeper by almost an order of magnitude. This project will look into various aspects of S-SMR and contribute to cutting-edge research with high prospects of applicable results within a team of highly motivated and talented students.

[MaRS-2024-02]
Blockchain, blockchain, blockchain
Contact: Prof. Fernando Pedone
Blockchain has gained much traction in recent years. From a topic restricted to specialized circles, it has made it to the general press with daily headlines, including many scandals. Beyond the hype, blockchain fostered the development of sophisticated distributed algorithms and models. And many interesting issues remain unaddressed. In this project, the student will team up with a group of talented researchers from USI and a leading blockchain enterprise to help advance the state of the art in the field.

[MaRS-2024-03]
Towards Efficient Dataset Reduction to Reduce the Costs of Mutation Testing for Deep Learning Systems
Contact: Prof. Paolo Tonella
Co-Supervisor: Dr. Nargiz Humbatova
Deep Learning (DL) has become an integral part of many groundbreaking projects and products that we use every day. As quality and safety remain the main concerns for developers and users of modern AI-based products, various techniques for assessing their quality are of increasing interest to the research community. Mutation testing is one such technique, in which errors in the form of small syntactic changes are deliberately introduced into the program under test to create a set of faulty programs, called mutants. The general principle underlying this approach is the assumption that faults used by mutation testing represent the mistakes that

programmers usually make. Mutation testing aims to assess the quality of a given test suite in terms of its capability to detect faults.

In traditional software systems the decision logic is often implemented by software developers in the form of source code. In contrast, the behaviour of a DL system is mostly determined by the training data set and the training program, i.e., these are the two major sources of defects for DL systems. There exists a mutation testing tool called DeepCrime [1], which is designed to perform mutation testing on DL systems and is based on real DLspecific faults. However, it injects faults into a DL system prior to the training following a realistic fault injection scenario and is therefore computationally expensive. There is another tool DeepMutation++[2], which is computationally cheap because it injects faults into an already trained model. Such changes are random and not very likely to occur in the real world. These mutations usually introduce small noise or changes to a randomly selected subset of weights or change the structure of an already trained DL model by adding/deleting its layers or replacing the activation function.

The aim of the project is to explore methods for effectively selecting subsets of data from the existing dataset to enable (1) smarter and faster DL-specific post-training mutations, and (2) reducing the cost of pre-training mutations of DL systems. In particular, in the first use case, the student will explore ways to group inputs based on some common features so that perturbing the behaviour of an already trained model can produce meaningful mutants. In the second application, the student will investigate different dataset reduction techniques, with the aim of reducing the training time by reducing the amount of training data fed to a model, while maintaining the same performance of the trained model or the same mutation test results. It is worth noting that efficient dataset selection/reduction can be useful for various tasks beyond mutation testing.

[1] Nargiz Humbatova, Gunel Jahangirova, and Paolo Tonella. 2021. DeepCrime: mutation testing of deep learning systems based on real faults.In Proceedings of the 30th ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA ’21)

[2] Hu, Qiang, Lei Ma, Xiaofei Xie, Bing Yu, Yang Liu, and Jianjun Zhao. Deepmutation++: A mutation testing framework for deep learning systems." In 2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE’19)

[MaRS-2024-04]
Enhancing Deep Learning Mutations via Weight-specific Fuzzing
Contact: Prof. Paolo Tonella
Co-supervisor: Dr. Jinhan Kim
Mutation testing in Deep Learning (DL) systems has been used to enhance the robustness and reliability of DL systems, by assessing their test sets using model variations, referred to as "mutants". One category of DL mutations is a post-training mutation, which perturbs the weights of trained models in order to subtly break the original model. This method provides a measure of the quality of a test set by identifying whether the test set can effectively discern the mutated model. The procedure of breaking the DL model employs several Mutation Operators (MOs) such as Gaussian Fuzzing (GF) -- an operator that targets a portion of weights to adjust their values based on a Gaussian distribution. However, this approach neglects the intrinsic properties of the weights. The impacts of MO should differ across the

weights, neurons, layers, and model structure, taking into account local characteristics. Therefore, this project intends to develop a new weight-specific MO that accommodates the distributions of weights derived from multiple original models. The student will develop this new mutation strategy, which involves two main steps. The initial phase requires understanding the state-of-the-art post-training mutation technique, using its source code as a base to build the new weight-specific MO. The next step involves assessing the performance of this new MO across different dataset/models to determine the sensitivity and killability of its mutants.

[MaRS-2024-05]
Test Input Prioritization for Autonomous Driving Systems
Contact: Prof. Paolo Tonella
Co-Supervisor: Tahereh Zohdinasab
Autonomous driving systems have garnered significant attention in recent years, with rapid advancements in technology and the increasing demand for self-driving vehicles. Ensuring the safety and reliability of these systems is of paramount importance before their deployment on public roads. Thorough testing is crucial to identify potential failures and assess the system's behavior under various scenarios. However, exhaustive testing of autonomous driving systems is resource-intensive, time-consuming and difficult due to the huge test input space to cover and the high labeling cost. Consequently, prioritizing test inputs and labeling only 'high quality' ones, i.e., tests more likely to trigger a failure, is necessary for test cost reduction.

Multiple test prioritization techniques have been proposed by Software Engineering researchers. These techniques are mainly evaluated by measuring the number of failures exposed by the prioritized tests. Simply counting the number of failures produces a distorted view on the quality of the prioritized test inputs since they may be extremely similar (or even identical) inputs that expose the same problem of the deep learning component. Ideally, the prioritized test set should contain very diverse tests. Therefore, defining and measuring test uniqueness is extremely important for test prioritization when considering inputs for DL systems.

This master thesis proposal aims to address the challenge of test input prioritization for autonomous driving systems, with the goal of enhancing testing efficiency while preserving diversity. The proposed research seeks to develop a method that can effectively prioritize the test inputs, enabling the identification of critical scenarios that pose higher risks or require specific attention during testing and to use a novel metric that takes into account also test uniqueness to evaluate and compare the existing test prioritization techniques.

The research methodology will involve the following key steps:

1) Literature Review: Conduct a comprehensive review of existing literature on autonomous driving systems, and test input prioritization techniques. Identify the limitations and gaps in current approaches.
2) Test Input Prioritization Framework: Design and implement a prioritization framework that incorporates various factors, such as safety concerns and diversity. The framework should efficiently rank the test inputs based on their importance and potential impact on system performance.

3) Performance Metrics and Evaluation: Define appropriate performance metrics to evaluate the effectiveness of the proposed prioritization method. Compare the prioritized test inputs with state of the art approaches in terms of the number of diverse critical scenarios detected and testing efficiency.

The expected outcome of this research is a novel test input prioritization method that enhances test diversity for autonomous driving systems. The proposed framework will aid researchers and developers in allocating testing resources effectively by focusing on critical scenarios and reducing redundant testing efforts. Furthermore, this research contributes to the broader field of autonomous driving by providing insights into the challenges and opportunities in testing and validating these complex systems.

[MaRS-2024-06]
Fast updates of the Constraint Delaunay Triangulation and other tree Voronoi structures
Contact: Prof. Evanthia Papadopoulou
The constrained Delaunay triangulation (CDT) is a variant of the well-known Delaunay triangulation in which specified edges, sometimes called segments, are constrained to appear. These constraints have many uses, such as representing boundaries of non-convex objects, supporting better interpolation of discontinuous functions, and aiding the enforcement of boundary conditions in finite element meshes [1]. The CDT of a point-set is as close to being a Delaunay triangulation as possible subject to those constraints. The project will develop randomized linear-time algorithms to update a constraint Delaunay triangulation when a new segment constraint is inserted. This will lead to a fast incremental construction of the constraint Delaunay triangulation. CDTs are used, among others, in Delaunay refinement methods for mesh generation. The project is a concrete example of a new approach to compute tree and forest Voronoi diagrams in linear expected time.

The proposed algorithms correspond to recent research results based on the following two papers.

[1] Fast segment insertion and incremental construction of constrained Delaunay triangulations, Jonathan Richard Shewchuk and Brielin C. Brown, Computational Geometry: Theory and Applications, 2015

[2] Abstract Voronoi-like Graphs: Extending Delaunay's Theorem and Applications, Evanthia Papadopoulou, SoCG 2022 https://drops.dagstuhl.de/storage/00lipics/lipics-vol258socg2023/LIPIcs.SoCG.2023.52/LIPIcs.SoCG.2023.52.pdf

[MaRS-2024-07]
Partially Homomorphic Encryption for Stream Processing Frameworks
Contact: Prof. Patrick Eugster
Co-Supervisor: Dr. Pavel Chuprikov
Due to the rapid spread of IoT, billions of devices are expected to continuously collect and process sensitive data. Because of limited computational and storage capacity available on IoT devices, the current de facto approach is to send the gathered data to the cloud for computation. Unfortunately, using public (untrusted) cloud infrastructures for processing

continuous queries including on sensitive data leads to concerns over data confidentiality. An attractive approach to preserving the confidentiality of continuous query processing while utilizing public clouds is through the use of partially homomorphic encryption (PHE). PHE allows computations over encrypted data, without revealing plaintext values. Recently, we developed a set of symmetric cryptosystems that retain the homomorphic expressiveness of previous asymmetric cryptosystems while being more performant (https://github.com/ssavvides/symmetria). The goal of this project is to apply these existing symmetric PHE schemes to allow confidentiality-preserving continuous queries, using the Apache Storm stream processing framework into which we previously already integrated several asymmetric PHE schemes. The student will take part in rewriting select streaming queries to use symmetric PHE, designing efficient serialization and deserialization strategies, and introducing ciphertext handling nodes into Storm's topologies.
Microbenchmarking may be further used to guide the design of resource usage.

[MaRS-2024-08]
Cost-based Mechanism Selection for Secure Cloud Computing
Contact: Prof. Patrick Eugster
Co-Supervisor: Dr. Pavel Chuprikov
The value of what can be derived from customer data is being increasingly recognized by many industries, making information the new currency. With the amount of data generated growing exponentially, many organizations have turned to the cloud in their search for costeffective information processing. This leads to security concerns of processing sensitive data using third-party untrusted computational resources. While there are both software (homomorphic encryption) and hardware (secure enclaves) techniques with the potential to perform such processing without leaking information, they have their own constraints and overheads, so that there is no single universal solution. We have designed a system, called Hydra, that supports a multitude of security mechanisms nicely decoupling privacy policies from queries. To guarantee compliance with a chosen privacy policy, Hydra introduces a domain specific language (DSL). Hydra is integrated with the Apache Spark streaming processor to provide users with the familiar query abstraction. The current limitation of Hydra is that the choice of the security mechanism is hardcoded, while our DSL is able to check compliance of any such choice. The student will be working on query execution heuristics in Hydra (which uses standard Spark SQL extension points), based on empirical performance measurements for different security mechanisms combined with Spark execution metrics.

[MaRS-2024-09]
Language-based Policy Checking for Secure Computing
Contact: Prof. Patrick Eugster
Co-supervisor: Dr. Pavel Chuprikov
The value of what can be derived from customer data is being increasingly recognized by many industries, making information the new currency. With the amount of data generated growing exponentially, many organizations have turned to the cloud in their search for costeffective information processing. This leads to security concerns of processing sensitive data using third-party untrusted computational resources. While there are both software (homomorphic encryption) and hardware (secure enclaves) techniques with the potential to perform such processing without leaking information, they have their own constraints and overheads, so that there is no single universal solution. We have designed a system, called

Hydra, that supports a multitude of security mechanisms nicely decoupling privacy policies from queries. A security policy checker in Hydra checks compliance of queries with the security policy of interest — in particular, ensuring absence of insecure information flows. Hydra is based on Apache Spark — one of the most active open source projects boasting contributions from over 1200 developers spread across 300 companies, making it a unique code base to learn and experience the length and breadth of system design principles. As part of the project, the student will integrate, enrich, and optimize the security policy checker with SparkSQL — Hydra's query processing pipeline.

[MaRS-2024-10]
Compile-time Verification of Fault-tolerant Distributed Systems
Contact: Prof. Patrick Eugster
Co-Supervisor: Dr. Pavel Chuprikov
Software defects cost our IT-centered society exorbitant amounts of money. To make matters worse, driven by the advent of paradigms such as cloud computing and blockchains, software has been becoming increasingly distributed, i.e., its execution spans many processes. Besides having to avoid "conventional" intrinsic defects in the actual software, programmers now have to cater for partial failures, e.g., the possibility that certain processes or hosts fail while others continue to operate. Catering for these requires complex protocols, making implementation error-prone. Traditional "full depth" verification of programs involve lengthy verification processes requiring much manual effort and expert knowledge and are thus easily left out of the loop. The goal of this project is to apply and improve a prototype fault-tolerant event-driven programming framework that allows verification of component interaction as part of compilation. Our target are distributed middleware systems following earlier experiences applying it to the cluster manager component of the Apache Spark system. Along the way, extensions and practical additions to the respective domainspecific language may be investigated together with runtime optimizations, and performance evaluation conducted to demonstrate low overheads.

[MaRS-2024-11]
Portable Programmer-agnostic use of Trusted Hardware
Contact: Prof. Patrick Eugster
Co-Supervisor: Dr. Pavel Chuprikov
Malware and other attempts of tampering with computer software remain a dominant challenge to computer security. While several trusted execution environments (TEEs) allow programs to be shielded from attacks (e.g., Intel SGX, ARM Trustzone, AWS Nitro), leveraging these requires expert knowledge in security and the respective TEE, in addition to deep understanding of the corresponding programs. Even without considering the performance characteristics of different TEEs, TEE-based programs are not portable across TEEs of different vendors due to different APIs and functionalities proposed. We thus propose to use TEEs in combination with program anomaly detection (AD). By creating models of programs and comparing these against executions at runtime, AD can be applied without modifications to software. By tracing appropriate features, AD can detect various attacks with low overhead. However, being implemented fully in software, existing AD solutions have a fundamental flaw – their own mechanisms for tracing executions and comparing them to models are not protected from tampering. The goal of this project is thus to implement an AD monitor in a secure manner leveraging a TEE and other hardware features. This generic implementation

of the monitor is independent of any monitored program thus supporting portability across TEEs of different vendors.

[MaRS-2024-12]
Unbounded Model Checking for TLA+
Contact: Prof. Patrick Eugster
Co-Supervisor: Dr. Rodrigo Otoni
Correctness of distributed systems is of critical importance. One good way to establish guarantees about the behavior of such systems is through the writing and reasoning of TLA+ specifications, with this approach having been used in practice for a number of years. One important limitation of existing automated reasoning support for TLA+ is that it is based on bounded techniques, weakening the guarantees given to users and potentially hiding vulnerabilities. The goal of this project is to address this problem via a novel unbounded model checking approach for TLA+, based on the solving of constrained Horn clauses (CHC). CHC-based automated reasoning is a growing field, in which successful results have been achieved already say in the context of blockchain verification, but whose potential has not yet been fully tapped. The goal of this project is to extend the open-source symbolic model checker Apalache to improve the scalability of bounded TLA+ verification. Developing a CHC-based reasoning engine for Apalache and exploiting recent advances in CHC solving will enable efficient unbounded verification of TLA+ specifications.

[MaRS-2024-13]
Formal Modeling of Probabilistic Quantum Network Policies
Contact: Prof. Patrick Eugster
Co-Supervisors: Dr. Anita Buckley, Dr. Pavel Chuprikov
Quantum computing, communication, and sensing technologies offer fundamentally new ways for information processing. The objective of quantum communication is to transmit quantum states, which may be entangled, causing stronger correlations. The no-cloning theorem (i.e., qubits cannot be copied) makes quantum communication inherently secure, leading to several novel applications. The distribution of entangled qubits (Bell pairs) between distant end-nodes will be the main task of the quantum Internet of the future, and the main challenge will be scaling. We are developing QNetKAT (Quantum NetKAT), a language and logic for dealing with and reasoning about quantum networks. QNetKAT has primitives for creating and transmitting Bell pairs, together with parallel and sequential composition operators, and offers a simple way for expressing quantum network policies. In the course of this project the student will get familiar with the components of quantum networks and protocols for long distance entanglement distribution. Decoherence, losses, and noise-errors cause stochastic behavior of quantum operations. The goal of this project is to develop the QNetKAT language with probabilistic semantics. The main tasks will consist in extending the language with new primitives for expressing probabilistic behaviors, and implementing these in the NetSquid quantum network simulation platform (https://netsquid.org/) using Python.

[MaRS-2024-14]
Automatic Feedback for PyTamaro Web Activities
Contact: Prof. Matthias Hauswirth
Co-Supervisor: Luca Chiodini

Over the last two years, the LuCE research group has been developing PyTamaro, an educational Python library to learn programming using graphics.
To ease the adoption of the library, the group has also developed an online platform, PyTamaro Web (https://pytamaro.si.usi.ch/), enabling learners to program with PyTamaro directly in their browser, without requiring any installation. The platform now contains more than 100 activities, which interleave explanations and code cells in which to write Python programs (in the style of a Jupyter notebook). The platform also contains a dozen of curricula that guide the learners through those activities.
PyTamaro Web is being actively developed and is used by hundreds of users per day, both in Swiss high schools and worldwide. It also hosts a curriculum part of the international Hour of Code initiative.

While working through the programming activities, learners can submit their code and see the output it produces, but so far they receive no indication about whether their solution (i.e., the produced graphic) is correct. The goal of this project is to extend the capabilities of PyTamaro Web to automatically check the correctness of the student's solution and automatically produce feedback, to help in the learning process.

Experience with Python and web development (React / TypeScript) is beneficial. If you are interested in programming languages and in research on teaching and assessing the understanding of programming, and eager to see your efforts having an immediate impact on thousands of students learning to program, contact us to learn more!

[MaRS-2024-15]
Better Documentation for Python Libraries
Contact: Prof. Matthias Hauswirth
Co-Supervisor: Luca Chiodini
Over the last two years, the LuCE research group has been developing PyTamaro, an educational Python library to learn programming using graphics.
To ease the adoption of the library, the group has also developed an online platform, PyTamaro Web (https://pytamaro.si.usi.ch/), enabling learners to program with PyTamaro directly in their browser, without requiring any installation. The platform now contains more than 100 activities, which interleave explanations and code cells in which to write Python programs (in the style of a Jupyter notebook). The platform also contains a dozen of curricula that guide the learners through those activities.
PyTamaro Web is being actively developed and is used by hundreds of users per day, both in Swiss high schools and worldwide. It also hosts a curriculum part of the international Hour of Code initiative.

A key aspect in working with any library, PyTamaro included, is being able to browse its documentation to quickly retrieve the piece of information needed (e.g., "which parameters does that function take? what is its return type?"). There are a number of standard tools to produce documentation (e.g., Sphinx). But despite the popularity of Python as a programming language, the documentation resulting from these tools can be extremely hard to navigate for beginner learners (and sometimes for proficient programmers as well!). The goal of this project is to develop a better format for the documentation of a library that takes into account the learners' needs at the various stages of the learning process. We would like to integrate this

into the PyTamaro Web platform for documenting PyTamaro, with possible extensions to other core parts of the Python standard library.

Experience with Python and web development (React / TypeScript) is beneficial. If you are interested in programming languages and in research on teaching and assessing the understanding of programming, and eager to see your efforts having an immediate impact on thousands of students learning to program, contact us to learn more!

[MaRS-2024-16]
Deep Learning of Diversification Processes: Simulation and Inference Across Disciplines
Contact: Prof. Ernst Wit
Co-supervisor: Dr. Francisco Richter Mendoza
Species diversification models, such as the birth-death model, provide crucial insights into the mechanisms driving species evolution and adaptation. These models are instrumental in explaining the emergence and extinction rates of species over time. Understanding these models is key to developing effective simulation frameworks, which can then be used to train deep neural networks. The training of neural networks with data generated from these models enables advanced statistical inference, uncovering patterns and relationships that are not immediately apparent. The potential of applying such models and inference techniques extends beyond biological systems. In social sciences, for instance, similar principles can be applied to study sociocultural evolution, economic dynamics, and population studies. This cross-disciplinary application showcases the versatility of the simulation framework combined with deep learning. The project will leverage the insights gained from species diversification models and other relevant literature, including key references such as [RJH+21] and [LLV+23], to develop a robust framework. This framework will not only model biological diversification but also simulate various social science scenarios. Subsequently, deep learning models, particularly neural networks, will be trained on these simulations for detailed statistical analysis and predictive modeling. The overarching goal is to bridge the gap between theoretical models of diversification and practical, data-driven applications in both natural and social sciences. By combining simulation, deep learning, and statistical inference, this project aims to provide a novel approach to understanding and predicting complex dynamic systems. This broad-spanning application of the proposed framework holds the potential to revolutionize how we approach problem-solving and decision-making in diverse fields.

References:
[LLV+23] I. Lajaaiti, S. Lambert, J. Voznica, H. Morlon, and F. Hartig. A comparison of deep learning architectures for inferring parameters of diversification models from extant phylogenies. bioRxiv, 03 2023.

[RJH+21] F. Richter, T. Janzen, H. Hildenbrandt, E. C. Wit, and R. S. Etienne. Detecting phylodiversity-dependent diversification with a general phylogenetic inference framework. bioRxiv, 07 2021.

[MaRS-2024-17]
XAI-Fed: Explainable AI for Federated Models in Wearable Sensing
Contact: Prof. Marc Langheinrich

Co-supervisors: Dr. Martin Gjoreski, Daniil Kirilenko

Federated learning and its combination with differential privacy is the latest technique for building privacy-aware machine-learning models. Its primary assumption – no data leaves the local data storage, has enabled its application in a variety of privacy-sensitive domains: mobile keyboard prediction, human mobility modeling based on GPS data, modeling from electronic health records, etc. Artificial Intelligence (AI) methods can bring significant and sustainable improvements to our lives. However, end-users must be able to understand those systems. Unfortunately, today's groundbreaking AI methods are black-boxed (i.e., the decision model and the process are not understandable). The increased complexity of AI algorithms has made previous eXplainable AI (XAI) tools unsuitable, including the fact that most of the XAI solutions are not designed to operate under privacy constraints. This project will investigate XAI techniques compatible with privacy-aware approaches (e.g., federated learning). The focus will be on counterfactual explainers [55] for wearable sensing data. Specific project tasks are: (i) Analyze XAI tools that can operate under privacy constraints, focusing on counterfactuals; (ii) Pre-process one dataset from wearable sensing systems. Example datasets include emotion recognition, activity recognition and energy expenditure estimation; (iii) Develop machine learning models for one of the datasets in step 2, and apply existing XAI tools on the models developed, including the method for generating counterfactual explanations, BayCon; (iv) Develop XAI tool for counterfactual explanations that can operate under privacy constraints.

[MaRS-2024-18]

Causal Attention for Concept Embedding Models

Contact: Prof. Marc Langheinrich

Co-supervisor: Dr. Pietro Barbiero

Concept Bottleneck Models (CBMs) represent a significant advancement in interpretable machine learning, functioning by constraining predictions to pass through an intermediate layer of human-understandable concepts. However, these models often face a trade-off between accuracy and interpretability. To address this, Concept Embedding Models (CEMs) have been developed, which use high-dimensional representations of concepts to maintain interpretability while enhancing model accuracy. Despite their progress, a key limitation in CEMs is the absence of causal mechanisms in task predictors, which currently rely on simple linear layers or basic neural models without considering the causal relevance of concepts to specific task labels.

The primary goal of this project is to innovate within the field of CEMs by designing a causal attention mechanism for task predictors. This mechanism aims to selectively emphasize the most relevant concepts for a given task, embedding a layer of causal understanding into the decision-making process. The project seeks to blend the interpretability of CBMs with the accuracy of high-dimensional concept representations, ultimately creating a model that not only performs well but also aligns with human reasoning and intervention strategies. This endeavor stands to significantly enhance both the efficacy and transparency of CEMs in practical applications.

[MaRS-2024-19]

Out-Of-Distribution Generalization in Concept Bottleneck Models via Latent Active Learning

Contact: Prof. Marc Langheinrich

Co-supervisor: Dr. Pietro Barbiero

Concept Bottleneck Models (CBMs) have become a cornerstone in interpretable machine learning, primarily due to their structure that forces predictions through an interpretable layer of concepts. This design significantly aids in both interpretability and targeted interventions. On the other hand, Active Learning, a paradigm where the model actively queries specific data points to label, optimizes learning efficiency, particularly in scenarios with limited labeled data. It becomes especially critical in handling Out-of-Distribution (OOD) data, which refers to data that differ significantly from the training distribution and often pose a challenge for models trained on specific datasets. Integrating active learning with generative CBMs can be particularly fruitful, as it enables the exploration and better understanding of the latent concept space, especially in OOD contexts.

The project aims to leverage active learning strategies to improve OOD generalization in CBMs. To this aim, the project seeks to navigate the latent concept space of generative CBMs, sample OOD embeddings from this space, and provide concept supervisions using an active approach. This approach will not only facilitate the model's adaptation to OOD scenarios but also enhance its overall interpretability and effectiveness in real-world applications where data distributions can significantly vary.

[MaRS-2024-20]

Interpretable Concept-based Semi-factuals

Contact: Prof. Marc Langheinrich

Co-supervisor: Dr. Pietro Barbiero

Concept Bottleneck Models (CBMs) have redefined the landscape of interpretable machine learning by ensuring that predictions are made through an explicit layer of human-understandable concepts. This bottleneck structure allows for direct interventions at the concept level, making CBMs invaluable for applications requiring transparency and explicability. An emerging area of interest in this domain is the generation of semi-factuals, which are plausible changes in concept labels which do not alter the downstream task prediction of the model. Semi-factuals serve as powerful tools in understanding model decisions by illustrating how changing some concept labels may not alter the final task prediction.

The goal of this project is to design and implement generative CBMs that can learn a latent concept space capable of modeling and generating concept-based semi-factuals. By doing so, we aim to enrich the interpretability of CBMs, providing users with meaningful insights about the boundaries and implications of potential interventions. The project focuses on enabling these models to sample from a distribution of semi-factuals at inference time, thereby offering a nuanced understanding of the model's decision-making process and the realistic scope of influencing these decisions.

[MaRS-2024-21]

Tabular deep concept reasoning

Contact: Prof. Marc Langheinrich

Co-supervisor: Dr. Pietro Barbiero

Tabular data, characterized by their structured format in rows and columns, are ubiquitous across various industries and scientific fields. Their applicability ranges from finance and healthcare to retail and beyond, making them one of the most common data types in machine learning applications. The ability to derive meaningful insights from tabular data is crucial, yet

challenging, due to the complexity and diversity of the data. In this context, Deep Concept Reasoner (DCR) emerges as a novel neural-symbolic approach that stands out for its ability to automatically generate interpretable probabilistic programs from predicted concepts. This approach bridges the gap between deep learning's predictive power and the interpretability of symbolic reasoning.
The aim of this project is to conduct a comprehensive evaluation of the Deep Concept Reasoner (DCR) in comparison to existing state-of-the-art machine learning models specialized in handling tabular data. The focus will be on two primary metrics: accuracy and interpretability. By assessing DCR's performance in these areas, the project seeks to establish its efficacy and potential as a tool for both accurate prediction and meaningful interpretation in applications dealing with tabular data. This evaluation will contribute to the understanding of neural-symbolic models' roles in practical machine learning tasks, particularly in scenarios where both high accuracy and clear interpretability are essential.

[MaRS-2024-22]
Development of a Spectral-Temporal Transformer for Enhanced Biomedical Signal Analysis
Contact: Prof. Marc Langheinrich
Co-supervisors: Dario Fenoglio, Dr. Martin Gjoreski
Transformers have become the leading machine learning models in various fields due to their singular ability to remember long-term dependencies and identify meaningful correlations for predictions. Surpassing traditional models like convolutional neural networks (CNNs), Long Short-Term Memory networks (LSTMs), and Recurrent Neural Networks (RNNs), they excel in sequence-based tasks, including natural language processing and time-series analysis [2]. This superiority is largely attributed to the attention mechanism, which allows the network to discern dependencies of each sequence element in relation to others.
Despite these advancements, the application of Transformers in biomedical signal analysis, such as in electrocardiograms (ECG) and electroencephalograms (EEG), is still under-explored. These signals conceal considerable diagnostic information within their spectral domain, essential for accurate medical assessments. Current deep learning networks, like STResNet, utilize Fourier transformation to capture critical frequency domain information from these signals.
This project proposes to develop a Spectral-Temporal Transformer model, inspired by STResNet's frequency analysis and transformers' long-term memory capabilities. The model aims to integrate spectral and temporal features in biomedical signals using an attention mechanism.

[MaRS-2024-23]
Self-Supervised Federated Learning for Sensor Data in The Wild
Contact: Prof. Marc Langheinrich
Co-supervisors: Dario Fenoglio, Mohan Li
While sensors and deep learning have achieved remarkable success in fields like human activity recognition (HAR), the Internet of Vehicles, or healthcare, one of the biggest challenges still facing the community is the exploration of unrefined, unlabeled data. Billions of sensor data have been generated daily from the edge devices, but labeling them for training models is an extremely laborious and knowledge-demanding task. As a result, much of this data remains unused, even though it could potentially improve models. Additionally, privacy

concerns restrict access to user data, further limiting the size of datasets available for data-hunger tasks.
This project focuses on two innovative approaches to address these issues: Self-supervised learning (SL) and the recent Federated Learning (FL). The SL community has seen rapid growth with the introduction of encoders, as unlabeled data has proven to be a valuable source for representation learning. FL, recently introduced by Google, offers a distributed learning framework that protects privacy by keeping user data locally while maintaining strong model performance. This project aims to integrate these two approaches into Self-Supervised Federated Learning (SSFL) specifically for sensor data, to enhance representation capabilities for downstream tasks such as HAR.

[MaRS-2024-24]
Human Activity Recognition with Multi-modality and Multi-frequency Federated Learning
Contact: Prof. Marc Langheinrich
Co-supervisors: Mohan Li, Dario Fenoglio
The evolution of Human Activity Recognition (HAR) technologies, especially through wearable sensors like head-worn devices, has opened new avenues for advanced personal health and activity monitoring. Traditional HAR systems, often relying on centralized machine learning, face significant privacy and data security challenges. Federated Learning (FL) offers a promising solution by enabling decentralized, privacy-aware machine learning across multiple devices. This approach ensures no data leaves the local device, thus addressing privacy concerns. In federated environments, however, there's a need for multi-sensor models that can collaboratively train across diverse user devices, overcoming the inherent heterogeneity. While the HAR community is rapidly expanding, the scope of datasets explored remains limited. Current research predominantly focuses on mobile devices or smartwatches, which datasets may not adequately distinguish specific head-related activities, such as eating or talking. Instead, head-worn devices present a more suitable alternative for such activities. Inspired by this observation, our project proposes to approach HAR with a multi-modality and multi-frequency strategy. The multi-frequency aspect allows the framework to adapt to various environments, including low battery scenarios, potentially reducing computing costs while maintaining high accuracy.

[MaRS-2024-25]
An integrated platform for workplace human sensing
Contact: Prof. Marc Langheinrich
Co-supervisors: Mohan Li, Pietro Barbiero
The relationship between employee productivity and job satisfaction has been an intriguing topic for decades. The benefits of maintaining a high efficiency at work are multi-folded: boost the self-confidence of employees which further promotes their productivity; keep a good work-life balance; create a satisfying work condition and relieve stress from working, etc. Feeling and being more productive are way more important than extending work time and exhausting yourself out, which is the trigger to mental illness such as depression, sleep disorder, even suicides under high pressure. However, quantitatively evaluating the productivity of workers is not an easy task. Even though sensors industry has been growing fast in recent decades, sensing productivity and satisfaction at work, despite in urgent need from each one of us, remains unsolved with a handy approach.

This project aims to address the relationship between productivity, job satisfaction, and well-being in the modern workplace, we initiate our first endeavor to build an integrated platform to collect sensor data among employees in workplace.

[MaRS-2024-26]
Personalization for stress and mood recognition in diverse user group
Contact: Prof. Silvia Santini
Co-supervisor: Lidia Alecci
Mood has a significant impact on how people behave, think, and act. Psychological and social science research highlights that physiological aspects, including mood and stress, vary between individuals. Consequently, a universal stress prediction model is ineffective due to these variations, as what works seamlessly for one person may yield inaccurate or inconsistent results for another. However, creating individual models for each person lacks scalability. To address this, the student will utilize an existing large dataset and apply clustering techniques to identify similarities among users. The aim is to develop a machine learning model optimized for the specific characteristics and behavior patterns within these user groups, providing a balanced and scalable solution for personalized stress prediction across a diverse population.

[MaRS-2024-27]
Protect privacy in wearable devices using data anonymization
Contact: Prof. Silvia Santini
Co-supervisor: Lidia Alecci
Wearables facilitate continuous data collection to monitor diverse human behaviors, covering activity, health, and stress. This personal data, including electrocardiogram, movements, and heart rate, is often accessible online for research. Despite the common practice of masking names with random identifiers, studies show that this is insufficient for user identity protection due to the subject-dependent nature of physiological data. This research utilizes existing data and models to implement and assess anonymization techniques such as noise addition and synthetic data generation. The objective is to determine the extent of effective user identity protection while minimizing disruption to human behavior prediction.

[MaRS-2024-28]
A Novel Approach Using Bilateral Data Fusion for EDA Data Classification
Contact: Prof. Silvia Santini
Co-supervisor: Leonardo Alchieri
This thesis addresses the impact of lateralization on Electrodermal Activity (EDA) sensors in wearable devices. Lateralization, influenced by brain hemisphere activation, affects the accuracy of EDA readings based on the device's placement on a specific body side. Despite recent studies highlighting this issue, there is limited exploration of the potential benefits of using EDA devices on both sides simultaneously. The research aims to fill this gap by investigating how leveraging data from both sides concurrently can enhance classifier accuracy through machine learning. The focus is on datasets in the lab, with implications for medical-grade applications affected by lateralization. Success in demonstrating improved accuracy may revolutionize the field, particularly in sensitive medical tasks, offering more reliable predictions for tasks impacted by lateralization.

[MaRS-2024-29]
Uncertainty-aware Deep Learning in digital healthcare
Contact: Prof. Silvia Santini
Co-supervisor: Leonardo Alchieri
In this thesis, the objective is to investigate the use of Monte Carlo Dropout, a Bayesian Deep Learning technique, to make uncertainty prediction on Neural Network outputs when dealing with physiological data. In particular, the student will investigate the creation of a system using Early Exit to adjust the computational power based on the uncertainty prediction.
The student will leverage state-of-the-art Deep Learning models to make predictions on publicly available datasets for health applications, e.g., ECG data from healthy and unhealthy individuals.

[MaRS-2024-30]
Empathetic Virtual Agents using Large Language Models
Contact: Prof. Silvia Santini
Co-supervisor: Nouran Abdalazim
The rapid advancement of pervasive systems and wearables has paved the way for personal informatics systems, e.g., personal assistants and chatbots. These systems can be deployed in different personal and professional settings. Such personal assistants aim at personalizing the user ‘s experience by taking into consideration the user’s affection status and respond based on the user’s emotions and mental state status.
In this project, we aim at developing an empathetic personal assistant tool that provides an affection-aware user experience. The proposed tool integrates users’ affection from wearables with large language models, e.g., ChatGPT.

[MaRS-2024-31]
Embodied Large Language Models for Personalized Meeting Summarization
Contact: Prof. Silvia Santini
Co-supervisor: Nouran Abdalazim
The rapid advancement of pervasive systems and wearables has paved the way for personal informatics systems, e.g., personal assistants and chatbots. Such systems gear towards enhancing users’ productivity in both work setting and personal life. In work settings, users leverage personal informatics systems in managing their tasks, tracking their progress and scheduling their meetings.
Meetings are crucial for communication, decision-making, and brainstorming. The effectiveness of meetings relies on participants’ ability to remember key topics, often using meeting minutes or notes as memory cues. However, this approach is time-consuming, demands high attention levels, and may lead to varied perceptions among participants, causing a lack of synchronization.
In this project, we aim at developing a personalized meeting summarization tool that aims at optimizing the meeting experience. The proposed tool integrates users’ affection from wearables with large language models, e.g., ChatGPT.