ΠΡΟΣΚΛΗΣΗ ΕΚΔΗΛΩΣΗΣ ΕΝΔΙΑΦΕΡΟΝΤΟΣ
ΓΙΑ ΧΟΡΗΓΗΣΗ ΥΠΟΤΡΟΦΙΩΝ

ΑΡΧΙΜΗΔΗΣ: Έρευνα στην Τεχνητή Νοημοσύνη, την Επιστήμη Δεδομένων και τους Αλγορίθμους

1. Ερευνητική Μονάδα «ΑΡΧΙΜΗΔΗΣ»

Αντικείμενο της Ερευνητικής Μονάδας «ΑΡΧΙΜΗΔΗΣ» είναι η βασική και η εφαρμοσμένη έρευνα στους τομείς της Τεχνητής Νοημοσύνης, της Επιστήμης Δεδομένων και των Αλγορίθμων. Στόχος είναι η Ε.Μ «ΑΡΧΙΜΗΔΗΣ» να καταστεί ένας κόμβος παγκόσμιας εμβέλειας για την έρευνα στις περιοχές αυτές, προσφέροντας ακαδημαϊκές και επαγγελματικές ευκαιρίες σε νέους ταλαντούχους επιστήμονες και, παράλληλα, δημιουργώντας μία γέφυρα μεταξύ της ελληνικής ερευνητικής κοινότητας και επιστημόνων που δρουν σε πανεπιστήμια του εξωτερικού.
Η Ε.Μ. «ΑΡΧΙΜΗΔΗΣ» λειτουργεί ως Μονάδα του Ερευνητικού Κέντρου «ΑΘΗΝΑ» (ΕΚ «ΑΘΗΝΑ») και χρηματοδοτείται από το Εθνικό Σχέδιο Ανάκαμψης και Ανθεκτικότητας «Ελλάδα 2.0» στο πλαίσιο του έργου "*Μονάδα «ΑΡΧΙΜΗΔΗΣ»: Έρευνα στην Τεχνητή Νοημοσύνη, την Επιστήμη Δεδομένων και τους Αλγορίθμους*». Περισσότερες πληροφορίες για το έργο βρίσκονται στον ιστότοπο https://www.archimedesai.gr/.

Επιστημονικός Διευθυντής της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ» είναι ο κ. Τίμος Σελλής, οι Κύριοι Ερευνητές του είναι ο Καθηγητής Κωνσταντίνος Δασκαλάκης (MIT) και ο Καθηγητής Χρίστος Παπαδημητρίου (Columbia University), και το Επιστημονικό του Συμβούλιο περιλαμβάνει τους Καθηγητές:

Anastasia Ailamaki (EPFL)
Leo Guibas (Stanford University)
Yannis Ioannidis (University of Athens)
Petros Maragos (National Technical University of Athens)
John Mylopoulos (University of Toronto)
Nikos Paragios (Ecole Centrale Paris)
Katia Sycara (CMU)
Sergios Theodoridis (Aalborg University)
John Tsitsiklis (MIT)
Mihalis Yannakakis (Columbia University)

Η έρευνα της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ» θα διεξάγεται σε ομάδες αποτελούμενες από έμπειρους ερευνητές και νέους επιστήμονες, κυρίως υποψήφιους διδάκτορες (PhD), αλλά και μεταδιδακτορικούς ερευνητές (Postdoc), καθώς και τελειόφοιτους προπτυχιακούς φοιτητές (interns).

2. Η προκήρυξη

Με την παρούσα «1^{η} Προκήρυξη Υποτροφιών ΑΡΧΙΜΗΔΗΣ» για Υποψήφιους/ες Διδάκτορες προσκαλούνται νέοι και νέες επιστήμονες προκειμένου να συμμετέχουν στο πλαίσιο της εκπόνησης Διδακτορικής Διατριβής, στις ερευνητικές δραστηριότητες της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ» υπό την επίβλεψη των Ερευνητών που συνεργάζονται με αυτήν.

Σε όσες και όσους επιλεγούν θα προσφερθεί χρηματοδότηση μέσω υποτροφίας για την εκπόνηση της διδακτορικής τους διατριβής σε Ανώτατα Εκπαιδευτικά Ιδρύματα (ΑΕΙ) στην Ελλάδα με την ταυτόχρονη διεξαγωγή υψηλού επιπέδου έρευνας στην Ε.Μ. «ΑΡΧΙΜΗΔΗΣ».
Η διάρκεια υποτροφίας είναι έως 48 μήνες.

Κάθε υπότροφος θα επιβλέπεται για την εκπόνηση της διατριβής του/της από Κύριους και Συνεργαζόμενους Ερευνητές της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ», που είναι οι παρακάτω:

Κύριοι Ερευνητές
Κωνσταντίνος Δασκαλάκης, MIT
Χρίστος Παπαδημητρίου, COLUMBIA UNIV

Συνεργαζόμενοι Ερευνητές
Γιώργος Αμανατίδης, UNIV OF ESSEX
Αντώνης Αναστασόπουλος, GEORGE MASON UNIV
Ίων Ανδρουτσόπουλος, ΟΠΑ
Μαρία Βακαλοπούλου, UNIV PARIS-SACLAY
Yang Cai, YALE UNIV
Δημήτρης Γαλάνης, ΕΡ. ΚΕΝΤΡΟ «ΑΘΗΝΑ»
Κώστας Δανιηλίδης, UNIV OF PENNSYLVANIA
Στέφανος Ζαφειρίου, IMPERIAL COLLEGE LONDON
Κωνσταντίνος Καραμανής, UNIV OF TEXAS AUSTIN
Νίκος Κομοντάκης, ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Βαγγέλης Μαρκάκης, ΟΠΑ
Στέλλα Μαρκαντωνάτου, ΕΡ. ΚΕΝΤΡΟ «ΑΘΗΝΑ»
Παναγιώτης Μερτικόπουλος, ΕΚΠΑ
Άρης Μουστάκας, ΕΚΠΑ
Άρης Παγουρτζής, ΕΜΠ
Ιωάννης Παναγάκης, ΕΚΠΑ
Ιωάννης Παναγέας, UNIV OF CALIFORNIA IRVINE
Μαρία Παπαδοπούλη, ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Γεώργιος Παπαναστασίου, UNIVERSITY OF ESSEX
Ευαγγελία Πιτουρά, ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
Χαρά Ποδηματά, MIT
Αγγελική Ράλλη, ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
Αλκμήνη Σγουρίτσα, ΟΠΑ
Κατερίνα Σωτηράκη, YALE UNIV
Χρήστος Τζάμος, ΕΚΠΑ
Παναγιώτης Τσαπάρας, ΠΑΝ/ΜΙΟ ΙΩΑΝΝΙΝΩΝ
Σωτήρης Τσαυτάρης, UNIVERSITY OF EDINBURGH
Δημήτρης Φωτάκης, ΕΜΠ
Γιώργος Χριστοδούλου, ΑΠΘ

Οι αποδέκτες υποτροφιών θα συνεργαστούν με τα στελέχη της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ» για την εξασφάλιση θέσης υποψηφίου διδάκτορα σε Ελληνικό ΑΕΙ. Η εγγραφή σε διδακτορικό πρόγραμμα αναμένεται να γίνει μέσα στο 2023.

Οι θεματικές περιοχές των προσφερόμενων υποτροφιών ενδεικτικά είναι:

- Causality and Fairness in Data Science and Machine Learning
- Game Theory / Optimization / Multi-Agent Learning
- Machine Learning and Computer Vision
- Machine Learning and Life Sciences
- Machine Learning and Natural Language Processing
- Machine Learning Foundations

Αναλυτική περιγραφή ενδεικτικών θεμάτων έρευνας βρίσκεται στο παράρτημα.

3. Το πλαίσιο της συνεργασίας

Κάθε υπότροφος θα υποστηριχθεί στην έρευνά του/της από την ομάδα των ερευνητών της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ», ενώ το περιβάλλον θα δώσει την δυνατότητα συμμετοχής σε διάφορες δραστηριότητες όπως ομιλίες, σεμινάρια, workshops, κλπ. που θα λαμβάνουν χώρα στο χώρο της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ». Επιπλέον:

3.1. Η μηνιαία αποζημίωση για το χρονικό διάστημα της υποτροφίας για διδακτορικές σπουδές ορίζεται στα 1.300€, ποσό το οποίο αντιστοιχεί στην αμοιβή του/της υποτρόφου με βάση τον εσωτερικό κανονισμό λειτουργίας του ΕΚ «ΑΘΗΝΑ». Για τη συνεργασία θα υπογράφεται σύμβαση υποτροφίας μεταξύ του/της ΥΔ και του ΕΚ «ΑΘΗΝΑ». Στους επιτυχόντες/επιτυχούσες θα προσφερθεί σύμβαση προπτυχιακής ή μεταπτυχιακής υποτροφίας (ανάλογα με την ακαδημαϊκή τους κατάσταση) με μηνιαίο ποσό 800€ ή 1.000€, αντίστοιχα, μέχρι να εγγραφούν σε πρόγραμμα διδακτορικών σπουδών.

3.2. Οι μετακινήσεις των ΥΔ που θα κριθούν απαραίτητες για την υλοποίηση του έργου τους, είτε για τη συμμετοχή σε συνέδρια είτε για άλλες ερευνητικές δραστηριότητες, θα καλύπτονται από το ΕΚ «ΑΘΗΝΑ».

3.3. Οι ΥΔ θα έχουν σαν βάση εργασίας τις εγκαταστάσεις της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ», συνεκτιμώντας και τις υποχρεώσεις τους που θα απορρέουν από την ιδιότητα του/της υποψήφιου/ας διδάκτορα.

4. Προϋποθέσεις συμμετοχής

Οι όροι και προϋποθέσεις για τη συμμετοχή στην παρούσα προκήρυξη είναι οι εξής:

- Να μην είναι ήδη κάτοχοι διδακτορικού διπλώματος σε οποιονδήποτε επιστημονικό τομέα.
- Να μην εκπονούν ήδη διδακτορική διατριβή σε Ελληνικό ΑΕΙ.

Αν κάποιος/α ενδιαφερόμενος/η δεν ικανοποιεί τις παρακάτω προϋποθέσεις μπορεί να ανατρέξει στις **Συχνές Ερωτήσεις (FAQs)** της προκήρυξης για πιθανές εξαιρέσεις. Επίσης, ερωτήσεις για την προκήρυξη μπορούν να υποβάλλονται **εδώ** και αν είναι γενικού ενδιαφέροντος, οι απαντήσεις θα αναρτώνται στον ισότοπο με τις συχνές ερωτήσεις. Για διευκρινήσεις, μπορείτε να στείλετε μήνυμα στο liaison-archimedes@athenarc.gr.

Επιπλέον, οι όροι για την λήψη της υποτροφίας από επιτυχόντες/επιτυχούσες είναι:

- Να μη χρηματοδοτούνται για την προτεινόμενη διδακτορική έρευνα (για το σύνολο ή και για τμήμα αυτής) από οποιαδήποτε άλλη πηγή (δημόσια, ιδιωτική, ευρωπαϊκή, διεθνή) κατά τη διάρκεια της υποτροφίας.
- Να μη λαμβάνουν άλλη υποτροφία από οιανδήποτε πηγή κατά τη διάρκεια της υποτροφίας.
- Να μην απασχολούνται ως πλήρους απασχόλησης στο Δημόσιο ή στον Ιδιωτικό Τομέα.
- Να μη λαμβάνουν επίδομα ανεργίας στην Ελλάδα ή το εξωτερικό κατά τη διάρκεια της υποτροφίας.

5. Υποβολή Αιτήσεων και Διαδικασία Επιλογής

5.1. Η αίτηση υποβάλλεται ηλεκτρονικά **εδώ**. Αν δεν είστε εγγεγραμμένοι χρήστες στον ιστότοπο του ΕΛ.ΙΔ.ΕΚ, ακολουθείστε τις οδηγίες για να εγγραφείτε πριν υποβάλλετε την αίτηση. Αφού κάνετε είσοδο (login) στο σύστημα, οδηγίες για τη χρήση του συστήματος υποβολών θα βρείτε εδώ.

5.2. Οι αιτήσεις μπορούν να υποβληθούν από τις **17/02/2023** και η καταληκτική ημερομηνία υποβολής αιτήσεων είναι η **31/03/2023,** και ώρα **17:00.**

5.3. Κάθε υποψήφιος/α μπορεί να υποβάλλει αίτηση για **μία έως τρεις θεματικές περιοχές**. Προϋπόθεση για μία επιτυχημένη συνεργασία είναι οι αιτήσεις να υποβάλλονται αφού ο/η υποψήφιος/α βεβαιωθεί ότι η θέση που διεκδικεί είναι συμβατή με τα ερευνητικά του/της ενδιαφέροντα.

5.4. Η αξιολόγηση των αιτήσεων και η επιλογή θα γίνει από τον επιστημονικό διευθυντή και τους ερευνητές της Ε.Μ. «ΑΡΧΙΜΗΔΗΣ». Στην αξιολόγηση θα συνεκτιμώνται τα παρακάτω:

- Τίτλοι σπουδών
- Βιογραφικό σημείωμα
- Γνώση ξένων γλωσσών
- Δημοσιεύσεις, Συμμετοχή/Ανακοινώσεις σε Συνέδρια
- Διακρίσεις ή/και προηγούμενες υποτροφίες
- Ερευνητική δραστηριότητα στο αντίστοιχο επιστημονικό πεδίο
- Δήλωση κινήτρων (Έκθεση σκοπιμότητας)
- Συστατικές επιστολές

Ελλάδα 2.0
ΕΘΝΙΚΟ ΣΧΕΔΙΟ ΑΝΑΚΑΜΨΗΣ ΚΑΙ ΑΝΘΕΚΤΙΚΟΤΗΤΑΣ
Με τη χρηματοδότηση της Ευρωπαϊκής Ένωσης
NextGenerationEU

ΠΑΡΑΡΤΗΜΑ

Thematic Areas and Indicative Thesis Topics

Field Machine Learning Foundations

Title **Computationally Efficient Algorithms for Robust Multi-Class Classification with Noise**

The context

In recent years, machine learning has seen enormous success achieving superhuman performance in many tasks that were never possible before. This success has been enabled by an abundance of data available from various domains which can be used to train complex ML models. Yet, an important bottleneck in applying ML techniques for many applications is the sensitivity of the existing methods to noise. It has been repeatedly observed that a small amount of noise can significantly hinder the performance of the ML models as errors in the training data tend to propagate in the model's predictions. Learning under noisy data is a challenging task in both theory and practice. Noise is quite widespread in many tasks and can range from adversarial, e.g. from data poisoning attacks, to more benign, e.g. caused by human annotation or measurement errors. In practice, many methods have been proposed to account and correct for the noise. The state-of-the-art approaches are inspired by theory but only have weak guarantees given the difficulty in obtaining theoretical results in this space. The goal of this project is to bridge the gap between theory and practice by (i) expanding the theoretical frontiers of settings that are solvable computationally efficiently, and (ii) applying this theory to obtain principled training pipelines that are competitive with the best methods in practice. The focus will be on more mild and/or structured noise models that are practically relevant and, at the same time, allow for computationally efficient solutions.

The topic

The goal of this direction is to extend the results of binary classification under semi-random noise to the multiclass setting and obtain algorithms with provable robustness and efficiency guarantees. The project will consist of two stages, with the first one focusing mostly on the theoretical side and the second one aiming to obtain practical algorithms for real-world applications. The first stage of the project will develop novel algorithmic techniques for classification under noisy labels. There is a long line of recent work, focusing on semi-random noise models aiming to understand the intricacies of the problem. The research has led to many novel insights on effective techniques for dealing with label noise, yet almost in its entirety, the work has focused on the fundamental case of binary classification. With the new toolkit in hand, the goal of this project is to address the more complex case of multiclass classification which is the leading task in modern Machine Learning. The multiclass case presents new difficulties that were not present in the binary case, and all known algorithms for binary classification under semi-random noise cannot be extended for more than 2 classes. The goal is to investigate these issues in detail presenting novel algorithmic approaches for learning. An interesting case of the general problem is dealing with structured noise among the different classes of labels and allowing the algorithm to incorporate coarser and more combinatorial information about the true label of an example. Finally, an important application of robust classification is to develop robust algorithms for contextual bandits in reinforcement learning. The second stage focuses on analyzing and improving practical Machine Learning algorithms for noisy classification. The first stage will provide a novel theoretical framework for dealing with noise addressing important challenges in Machine Learning. The goal of the framework is to be widely applicable and practically relevant and at the same time tractable, namely solvable in polynomial time, and amenable to analysis. The second stage aims to drive this research forward by examining existing practical algorithms through the lens of the developed framework. An important direction is to provide performance guarantees for those algorithms or identify their potential shortcomings. In the latter case, the goal will be to extend the existing methods to incorporate the theoretical ideas from the first stage. The final objective is to develop novel iterative approaches that fit the standard Machine Learning pipeline and can be applied effectively to large datasets.

Background required

Machine learning principles, Probability Theory, Algorithms and Complexity

Desired qualifications

Python programming with experience in training Deep Neural Networks

Supervisors

Primary Supervisor: Christos Tzamos (National and Kapodistrian University of Athens, GR), http://tzamos.com

Title **Effective Methods for Data Cleaning and Reliable Machine Learning**

The context

In recent years, machine learning has seen enormous success achieving superhuman performance in many tasks that were never possible before. This success has been enabled by an abundance of data available from various domains which can be used to train complex ML models. Yet, an important bottleneck in applying ML techniques for many applications is the sensitivity of the existing methods to noise. It has been repeatedly observed that a small amount of noise can significantly hinder the performance of the ML models as errors in the training data tend to propagate in the model's predictions. Learning under noisy data is a challenging task in both theory and practice. Noise is quite widespread in many tasks and can range from adversarial, e.g. from data poisoning attacks, to more benign, e.g. caused by human annotation or measurement errors. In practice, many methods have been proposed to account and correct for the noise. The state-of-the-art approaches are inspired by theory but only have weak guarantees given the difficulty in obtaining theoretical results in this space. The goal of this project is to bridge the gap between theory and practice by (i) expanding the theoretical frontiers of settings that are solvable computationally efficiently, and (ii) applying this theory to obtain principled training pipelines that are competitive with the best methods in practice. The focus will be on more mild and/or structured noise models that are practically relevant and, at the same time, allow for computationally efficient solutions.

The topic

The goal of this direction is to examine and improve the reliability of existing machine learning algorithms under noisy data and develop effective methods to identify and correct the noisy inputs. The focus of this topic will be towards implementing reliable algorithms that work well in practice and measure their performance on popular benchmark datasets. The project will consist of two stages, with the first one focusing mostly on the theoretical side and the second one aiming to obtain practical algorithms for real-world applications. The first stage of the project will develop novel algorithmic techniques for classification under noisy labels. There is a long line of recent work, focusing on semi-random noise models aiming to understand the intricacies of the problem. The research has led to many novel insights on effective techniques for dealing with label noise, yet almost in its entirety, the work has focused on the fundamental case of binary classification. With the new toolkit in hand, the goal of this project is to address the more complex case of multiclass classification which is the leading task in modern Machine Learning. The multiclass case presents new difficulties that were not present in the binary case, and all known algorithms for binary classification under semi-random noise cannot be extended for more than 2 classes. The goal is to investigate these issues in detail presenting novel algorithmic approaches for learning. An interesting case of the general problem is dealing with structured noise among the different classes of labels and allowing the algorithm to incorporate coarser and more combinatorial information about the true label of an example. Finally, an important application of robust classification is to develop robust algorithms for contextual bandits in reinforcement learning. The second stage focuses on analyzing and improving practical Machine Learning algorithms for noisy classification. The first stage will provide a novel theoretical framework for dealing with noise addressing important challenges in Machine Learning. The goal of the framework is to be widely applicable and practically relevant and at the same time tractable, namely solvable in polynomial time, and amenable to analysis. The second stage aims to drive this research forward by examining existing practical algorithms through the lens of the developed framework. An important direction is to provide performance guarantees for those algorithms or identify their potential shortcomings. In the latter case, the goal will be to extend the existing methods to incorporate the theoretical ideas from the first stage. The final objective is to develop novel iterative approaches that fit the standard Machine Learning pipeline and can be applied effectively to large datasets.

Background required

Machine learning principles, Probability Theory, Algorithms and Complexity

Desired qualifications

Python programming with experience in training Deep Neural Networks

Supervisors

Primary Supervisor: Christos Tzamos (National and Kapodistrian University of Athens, GR), http://tzamos.com

Title **Safe and Reliable ML for Decision Making through Cryptographic Primitives**

The context

Complex automated systems using ML are being deployed to make consequential decisions for people's lives, like loan approval and probation decisions. Due to how consequential these decisions are, people are oftentimes incentivized to strategize with their data in an effort to obtain better decisions for themselves. In every ML system for decision making there are two key stakeholders: the organization making the decisions and the users/agents submitting their data for said decisions to be made. Deeply understanding both the organization and the user perspective is crucial for the safe and reliable deployment of ML for decision-making. On the one hand, the organizations want to make sure that their decisions are accurate, efficient, and robust, while their proprietary technology and data remain private. On the other hand, the users want to make sure that their data is not mistreated and that they have the opportunity for recourse through these automated systems.

The topic

The main objective of this PhD is to bridge the tension between the ML designer's and the individuals' perspective. We posit that cryptographic primitives are the indispensable tools that will allow us to achieve this goal. The work to be carried out will involve two threads. The first thread adopts the organization perspective, while having some of the user's considerations as constraints. The second thread adopts the user perspective, while having some of the organization's considerations as constraints. Fostering an inclusive environment is one of our core values. We seek candidates who will create a climate that helps attract and is inclusive of all students, including students from historically underrepresented groups and students who have overcome personal challenges. We strongly encourage women and underrepresented minorities to apply.

Background required

Mathematical Maturity, Analysis of Algorithms

Desired qualifications

Machine learning principles, Cryptography

Supervisors

Primary Supervisors: Chara Podimata (MIT, USA), https://www.charapodimata.com/

Title **Algorithms for Robust Reinforcement Learning**

The context

In recent years, machine learning has seen enormous success achieving superhuman performance in many tasks that were never possible before. This success has been enabled by an abundance of data available from various domains which can be used to train complex ML models. Yet, an important bottleneck in applying ML techniques for many applications is the sensitivity of the existing methods to noise. It has been repeatedly observed that a small amount of noise can significantly hinder the performance of the ML models as errors in the training data tend to propagate in the model's predictions. Learning under noisy data is a challenging task in both theory and practice. Noise is quite widespread in many tasks and can range from adversarial, e.g. from data poisoning attacks, to more benign, e.g. caused by human annotation or measurement errors. In practice, many methods have been proposed to account and correct for the noise. The state-of-the-art approaches are inspired by theory but only have weak guarantees given the difficulty in obtaining theoretical results in this space. The goal of this project is to bridge the gap between theory and practice by (i) expanding the theoretical frontiers of settings that are solvable computationally efficiently, and (ii) applying this theory to obtain principled training pipelines that are competitive with the best methods in practice. The focus will be on more mild and/or structured noise models that are practically relevant and, at the same time, allow for computationally efficient solutions.

The topic

The goal of this direction is to obtain provably robust and efficient algorithms for reinforcement learning. An important focus will be on designing noise tolerant bandit algorithms that work beyond the standard realizability assumptions required by existing algorithms. The project will consist of two stages, with the first one focusing mostly on the theoretical side and the second one aiming to obtain practical algorithms for real-world applications. The first stage of the project will develop novel algorithmic techniques for classification under noisy labels. There is a long line of recent work, focusing on semi-random noise models aiming to understand the intricacies of the problem. The research has led to many novel insights on effective techniques for dealing with label noise, yet almost in its entirety, the work has focused on the fundamental case of binary classification. With the new toolkit in hand, the goal of this project is to address the more complex case of multiclass classification which is the leading task in modern Machine Learning. The multiclass case presents new difficulties that were not present in the binary case, and all known algorithms for binary classification under semi-random noise cannot be extended for more than 2 classes. The goal is to investigate these issues in detail presenting novel algorithmic approaches for learning. An interesting case of the general problem is dealing with structured noise among the different classes of labels and allowing the algorithm to incorporate coarser and more combinatorial information about the true label of an example. Finally, an important application of robust classification is to develop robust algorithms for contextual bandits in reinforcement learning. The second stage focuses on analyzing and improving practical Machine Learning algorithms for noisy classification. The first stage will provide a novel theoretical framework for dealing with noise addressing important challenges in Machine Learning. The goal of the framework is to be widely applicable and practically relevant and at the same time tractable, namely solvable in polynomial time, and amenable to analysis. The second stage aims to drive this research forward by examining existing practical algorithms through the lens of the developed framework. An important direction is to provide performance guarantees for those algorithms or identify their potential shortcomings. In the latter case, the goal will be to extend the existing methods to incorporate the theoretical ideas from the first stage. The final objective is to develop novel iterative approaches that fit the standard Machine Learning pipeline and can be applied effectively to large datasets.

Background required

Machine learning principles, Probability Theory, Algorithms and Complexity

Desired qualifications

Python programming with experience in training Deep Neural Networks

Supervisors

Primary Supervisor: Christos Tzamos (National and Kapodistrian University of Athens, GR), http://tzamos.com

Title **Reinforcement Learning in Partially Observed Environments**

The context

Reinforcement Learning (RL) is a central problem at the intersection of optimization, control and machine learning, and represents a fundamental frontier in artificial intelligence for both theoretical and applied areas. At its heart, Reinforcement Learning is about sequential decision making in an unknown dynamic environment. Actions taken must balance between the immediate reward they bring, the potential information gain they represent, and the future advantages an action now may yield. Markovianity -- the basic assumption that past and future are independent, conditioned on the present state -- is a central assumption for RL, critical for most algorithms and analytical results. Yet in many important applications in robotics, autonomous driving, health-care, e-commerce and beyond, the agent has access to only partial observations of this state. In this case, optimal algorithms depend on the full history of observations and actions, thus making the problem considerably more challenging.

The topic

The main objective of this line of research is to develop algorithms and analysis techniques for important settings in RL where the decision-maker has access to only partial observations of the state. We are primarily interested in three settings: (A) novel and potentially adversarial noise models in the observation process, that may require developing new estimation techniques, whose near-optimality will depend on understanding and leveraging the dynamics of the overall process; (B) problems with latent variables, i.e., problems where we are unable to observe our environment fully, yet what we cannot observe changes slowly or not at all (for example, there only have partial (or no) genetic information about a patient we wish to treat, or we only know some characteristics of users in an online setting); (C) problems whose large scale forces us to work with a subset of the state or the observations, in the interest of computational tractability. This is a flexible project space, intended to accommodate several Ph.D. students, and with significant potential for expansion and collaboration.

Background required

Machine learning, Algorithms, Optimization, Linear Algebra, Probability

Desired qualifications

This project is primarily theory-oriented, and so is a good fit for students interested in theoretical research and contributions.

Supervisors

Primary Supervisor: Constantine Caramanis (UT Austin, USA), https://caramanis.github.io/)

Field Game Theory / Optimization / Multi-Agent Learning

Title **Algorithms for computing Nash equilibria in Markov Games**

The context

Reinforcement Learning (RL) has been a fundamental driver for recent advances in Artificial Intelligence, ranging from super-human performance in games including Go and Starcraft to decision making, autonomous driving and robotics. A core ingredient behind the success of single-agent RL systems, which are typically modelled as Markov Decision Processes (MDPs), is that an optimal stationary policy always exists and moreover, there exist efficient algorithms that provably converge to it. Recent progress has also been made towards understanding Multi-agent RL (MARL) settings with two players in which the reward of one player is the opposite of the reward of the other player, i.e., zero-sum stochastic games. However, in practice, a majority of the systems involve multi-agent interactions, going beyond single-agent and two-player zero-sum settings. In such cases, despite the notable empirical advancements, there is a lack of understanding of the theoretical convergence guarantees of existing MARL dynamics.

The topic

The aim of this PhD is to investigate MARL settings in which convergence to Nash policies - one of the most common solution concepts in MARL - can be provably guaranteed, despite the recent established negative results by Daskalakis et al. In other words, the main objective is to design and analyze Markov games with structure so that the problem of equilibrium computation becomes tractable: Such structural games include but are not limited to (1) classes of Markov potential games, (2) classes of team games, (3) constrained MARL settings. Although the PhD will be mainly theoretical, it will be of great interest to conduct empirical research on the above.

Background required

Algorithmic Game Theory, Learning in Games, Markov Decision Processes, Probability, Algorithms, Optimization

Desired qualifications

C, C++, Python (numpy) programming

Supervisors

Primary Supervisor: Ioannis Panageas, UC Irvine, https://panageas.github.io/

Title **Computational complexity of counting problems**

The context

Counting problems find extensive applications to a number of scientific fields, for example in statistical physics, learning theory, and computational social choice, among others. Despite significant progress in the field over the last 40+ years, important tantalizing questions remain unanswered, such as the approximability of counting the number of perfect matchings in general graphs. Progress in such fundamental questions is expected to have a remarkable impact on both the theory and practice of algorithms.

The topic

The main goal of the thesis is the classification of the computational difficulty of counting problems through methods from the areas of structural, descriptive, and parameterized complexity, as well as the development of new, more efficient algorithms for solving them. The development of algorithms will not be limited to classical deterministic approaches, but will also extend to probabilistic as well as approximation algorithms and schemes. In addition, novel approaches will be considered, such as fine-grained complexity, e.g. what is the best possible approximation in subquadratic time, and data driven algorithms, taking advantage of information on the input distribution that one might be able to extract by observing a number of instances. The ultimate goal is to take advantage of recent research in the field, and also to develop new techniques and ideas, in order to derive a complete picture of the difficulty and algorithms for such problems.

Background required

Algorithms and complexity, theory of computation

Desired qualifications

Counting problems, Markov chains, descriptive complexity

Supervisors

Primary Supervisor: Aris Pagourtzis (National Technical University of Athens, GR), https://www.ece.ntua.gr/en/staff/79

Title **Data-driven and learning-augmented mechanism design**

The context

Mechanism design is an important subfield of game theory with many success stories in practice. Auction mechanisms, in particular, range from high-stake governmental license auctions (among others, for telecommunications spectrum or carbon allowance), to procurement auctions for hiring sub-contractors, sponsored search auctions run by search engines like Google, and even to auctions run by single individuals in online marketplaces. While auctions have been extensively studied for decades, the landscape formed by modern large-scale applications, abundant data, and the emergence of machine learning as a prediction tool, entails revisiting the classic models of mechanism design and creates exciting new challenges. Can we use past data to find approximately optimal reserve prices for future auctions? What if we have access to noisy data or only a handful of samples? Could we design mechanisms with better properties based on ML-derived predictions of some parameters of the problem at hand?

The topic

The objective of this PhD will be to study mechanism design under various learning-based approaches for the informational assumptions the mechanism designer may use. Each of these approaches gives rise to a new research direction. The focus will be on auction mechanisms and the work to be carried out will involve two main topics: a) The design of mechanisms under data-driven paradigms, where the auctioneer can exploit access to data of past auctions. The goal here is to obtain new learning algorithms, applicable to expressive models of the bidders' preferences and identify strong guarantees regarding the volume of required samples and the level of achievable economic performance. Specifically for the setting where the samples arrive in an online fashion, the goal is to design mechanisms with vanishing regret. b) The study of learning-augmented scenarios, where one can utilize predictions on features of candidate solutions. The goal is to attain "best-of-both-worlds" mechanisms that have an improved performance when the prediction is precise, yet they maintain good worst-case guarantees when the prediction fails. Finding truthful, learning-augmented mechanisms is a largely unexplored territory in the literature.

Background required

Design and analysis of algorithms, Game theory, Optimization

Desired qualifications

BSc degree / Diploma in computer science, electrical engineering, operations research, or mathematics, MSc degree in computer science, electrical engineering, operations research, mathematics, or related topic, familiarity with programming in Python

Supervisors

Primary Supervisor: Georgios Amanatidis (University of Essex, UK), http://amanatidis.info/

Title **Data-driven approximation and online algorithms**

The context

Despite its apparent benefits and wide acceptance, worst-case analysis in algorithms is often simplistic and leads to conclusions that contradict common practical observations. In addition to few other factors, criticism to worst-case analysis has been fueled by the fact that modern algorithms typically operate on very large instances, which due to their size and generative process, are bound to exhibit patterns. Techniques from statistical and online learning are able to extrapolate such patterns and to successfully exploit them in order either to optimally determine crucial parameters of algorithms or to "predict" actions that would perform well on future unknown data. It is important to understand how (and how much) approximation and online algorithms can benefit from this approach. The goal is a new generation of algorithms that would exhibit simplicity and superior performance by getting adapted actively to the data on which they work. In the era of Big Data, Artificial Intelligence and Machine Learning, this is a key challenge of major scientific and practical importance for the algorithmic community.

The topic

The goal of the thesis is to develop novel techniques and approaches for the theoretical analysis of data-driven algorithms and to work towards a deeper understanding of how (and how much) approximation and online algorithms can benefit from a data-driven approach. The thesis will focus on fundamental problems in approximation and online algorithms, including constraint satisfaction, network design, metric facility location, clustering and scheduling. As potential approaches, the thesis will revisit known guarantees on the performance of data-driven algorithm from the viewpoint of universal learning, thus focusing on particular data distributions. Moreover, the thesis will aim to simplify and unify the design and analysis of non-adaptive and partially adaptive data-driven algorithms, by establishing a close connection between data-driven algorithm design and universal and a-priori approximation algorithms.

Background required
Algorithms and complexity, approximation and online algorithms, randomized algorithms

Desired qualifications
Some experience on the design and analysis of approximation, online and randomized algorithms

Supervisors
Primary Supervisor: Dimitris Fotakis (National Technical University of Athens, GR), https://www.ece.ntua.gr/en/staff/180

Title **Equilibrium Computation and Learning in Multi-Agent Systems**

The context
Machine Learning (ML) has witnessed tremendous success in various applications in the past decade. A particularly successful paradigm is to formulate the training of the ML system as minimizing a single loss function and solve it via some efficient optimization algorithm. Many emerging applications of ML start deviating from this single objective viewpoint and are better captured as a game between multiple intelligent agents/algorithms. The game could be explicit such as in markets, traffic routing, and multi-agent RL systems, or implicit such as in generative adversarial networks, adversarial examples, robust optimization, etc. Game theory provides the framework for us to think about the interaction between agents who each have their own loss function to minimize, but traditional game theory falls short in addressing the complex challenges posed by modern ML systems. This is because most ML applications involve high-dimensional and non-concave games, whereas game theory primarily deals with low-dimensional concave games.

The topic
The main objective of this PhD is to work towards a new theory for high-dimensional, non-concave games that is useful for ML applications. The focus is on defining meaningful solution concepts in such games and determining their computability. The research will encompass three key topics: (i) min-max optimization in normal form games, (ii) equilibrium computation and learning for multi-player stochastic games, and (iii) exploring new and alternative solution concepts.

Background required
Machine learning, Online learning, Game theory

Desired qualifications
Python programming, Optimization

Supervisors
Primary Supervisor: Yang Cai (Yale University, USA), http://www.cs.yale.edu/homes/cai/

Title **Learning, optimization and fairness aspects of multi agent systems**

The context
During the last years, we have experienced a fruitful interaction between optimization, machine learning and game theory. One of the main reasons is that several learning applications, including the training of certain deep learning models, essentially boil down to optimization problems over continuous action spaces, which in turn can be formulated as the problem of computing game-theoretic solutions (e.g., Nash equilibria) in appropriately defined games. This has revived the interest of the machine learning community both in games and in optimization. In particular, coming up with faster optimization algorithms can lead to significant improvements with a potentially positive effect on practical aspects as well. At the moment however, we are still missing a better understanding of the power and limitations of existing algorithms, that would also enlighten us towards developing new techniques for addressing yet unresolved questions.

The topic
The main objective of this PhD thesis is to contribute to the ongoing interplay between optimization theory and machine learning, with an emphasis on developing optimization approaches for the training of learning algorithms and more generally for multi-agent learning models. The work to be carried out is expected to focus primarily on the theoretical foundations of such approaches, followed by appropriate experimental or empirical performance evaluation. The thesis will evolve around the following main topics: (i) min-max optimization problems and their connections to computational learning theory (and especially to deep learning models); (ii)

first-order methods, both under deterministic and stochastic gradient information, for solving relevant optimization problems for training neural networks; (iii) game-theoretic solution concepts and algorithmic considerations, in the context of multi-agent learning.

Background required
Design and analysis of algorithms, machine learning principles, optimization theory, game theory

Desired qualifications
Undergraduate studies in computer science, electrical engineering or mathematics, M.Sc. degree in data science or related subject, familiarity with programming in Python

Supervisors
Primary Supervisor: Evangelos Markakis (Athens Univ of Economics and Business, GR), http://pages.cs.aueb.gr/~markakis/

Title **Learning-augmented approach to algorithmic mechanism design**

The context
Despite its apparent benefits and wide acceptance, worst-case analysis in algorithms is often simplistic and leads to conclusions that contradict common practical observations. In addition to few other factors, criticism to worst-case analysis has been fueled by the fact that modern algorithms typically operate on very large instances, which due to their size and generative process, are bound to exhibit patterns. Techniques from statistical and online learning are able to extrapolate such patterns and to successfully exploit them in order to optimally determine crucial parameters of truthful mechanisms (e.g., the item prices for auctions based on demand queries, or a near optimal clustering for multiple facility location games). It is important to understand how (and how much) truthful mechanism design, which is all about incentive compatible information elicitation, can benefit from this approach. The goal is a new generation of truthful mechanisms with near optimal performance guarantees, whose strong incentive compatibility properties would be facilitated by pieces of crucial information (possibly provided by a machine learning oracle) about the instance. In the era of Big Data, Artificial Intelligence and Machine Learning, this is a key challenge of major scientific and practical importance for the community of algorithmic mechanism design.

The topic
The goal of the thesis is to develop novel techniques and approaches for the design and analysis of learning-augmented truthful mechanisms and to work towards a deeper understanding of how (and how much) truthful mechanisms can benefit from a learning-augmented approach. The thesis will focus on fundamental problems in algorithmic mechanism design, including combinatorial auctions with submodular and subadditive bidders, procurement auctions, scheduling on self-interested machines, and facility location games. The thesis will aim to investigate how standard approaches in algorithmic mechanism design (e.g., monotone allocations, maximal-in-range mechanisms with VCG payments, item pricing combined with demand queries) could benefit from additional pieces of crucial information about the instance provided by a machine learning oracle. Another key objective is to better understand the incentives of the agents providing the data on which the machine learning oracle is trained.

Background required
Algorithms and complexity, algorithmic game theory, algorithmic mechanism design

Desired qualifications
Some experience on the design and analysis of truthful mechanisms

Supervisors
Primary Supervisor: Dimitris Fotakis (National Technical University of Athens, GR), https://www.ece.ntua.gr/en/staff/180

Title **Mechanism design augmented with learning**

The context
Why it is important: Mechanism design, a celebrated branch of Game Theory and Microeconomics, studies a special class of algorithms, called mechanisms, which are robust under selfish behaviour and produce a social outcome with a certain guaranteed quality. Unlike traditional algorithms that get their input from a single user, mechanisms solicit the input from different participants, in the form of preferences over the possible outputs (outcomes). The challenge stems from the fact that the actual preferences of the participants are private information, unknown to the algorithm. The participants are assumed to be utility maximisers that will provide

some input that suits their objective and may differ from their true preferences. A truthful mechanism provides incentives such that a truthful input is the best action for each participant. Unfortunately, truthfulness imposes dramatic limitations in algorithm design which are manifested in various forms, including information-theoretic (e.g., scheduling problems) and communication complexity aspects (e.g., combinatorial auctions). To alleviate both these issues, we plan to design mechanisms that use machine learning (ML) to exploit historical data.

The topic
The main objective of this PhD is to develop a robust theory for the design and analysis of mechanisms boosted by ML. We will deploy the newly developed model of algorithms with predictions which offers a perfect ground to combine the robust, but too pessimistic, worst-case analysis, with the impressive predictive power of ML. We will also use ML to produce good estimates of prior distributions. The aim is to replicate the great success of optimal mechanism design, but without the unrealistic assumption that the designer has access to (samples of) a prior distribution. The work to be carried out will involve 3 main questions/topics: (1) design and analysis of deterministic and randomized scheduling mechanisms with ML advice (2) design and analysis of combinatorial auctions with ML advice with low communication complexity (3) study the limitations of mechanism design with ML advice.

Background required
Algorithms and complexity

Desired qualifications
Algorithmic Game Theory, Approximation/Randomized/Online algorithms, Machine Learning

Supervisors
Primary Supervisor: Giorgos Christodoulou (Aristotle University of Thessaloniki, GR), https://sites.google.com/view/gchristo

Title **Optimization of deep generative models**

The context
Even though we are currently able to engineer machine learning systems that perform exceedingly well in specific tasks, the resulting models tend to have a narrow flight envelope, being brittle, data-hungry, and unable to adapt to tasks that deviate from their training configuration. This situation is exacerbated further when ML architectures interact with each other – e.g., as in the case of generative adversarial networks (GANs) or other generative / multi-agent systems. In that case, the training landscape becomes exponentially more intricate and can no longer be captured by the minimization of an empirical loss function over independent and identically distributed (i.i.d.) data. In turn, this has triggered a paradigm shift in the mathematical and conceptual foundations of ML model training, from "business as usual" minimization to min-max optimization and, more broadly, to equilibrium learning – i.e., learning an operating state which is unilaterally stable for all systems and agents involved.

The topic
The main goal of this PhD will be to design implementable algorithms capable of negotiating the fundamental obstructions faced by gradient methods in min-max problems (or, more generally, equilibrium problems). This will involve combining techniques from zeroth-order optimization and multi-armed bandits (stochastic or non-stochastic) with the analysis of standard gradient methods for neural network training. More specifically, the research involved will revolve around the following axes: (1) zeroth-order methods for min-max optimization with access to an approximate maximization oracle (and the verification of such oracles in deep learning); and/or (2) mixed-strategy reformulations of the problem as a bilinear game in higher dimensions and the use of continuous bandits for optimization. In both cases, the candidate will develop a deep understanding and contribute to the state-of-the-art of continuous optimization and bandit methods for generative models.

Background required
Mathematical fundamentals of optimization and machine learning

Desired qualifications
Expertise in game theory and multi-armed bandits will be a plus (but not required

Supervisors
Primary supervisor: Panayotis Mertikopoulos (CNRS, FR), http://polaris.imag.fr/panayotis.mertikopoulos

Title **Statistical Physics Inspired Methods for Loss Landscape Analysis in Machine Learning**

The context

Why it is important: Machine learning heavily depends on local training methods, which are ill-adapted in converging to the global optimum of the loss function at hand. Instead, if at all, they tend to converge to shallow critical (saddle) points. Additionally, the conditions for convergence to a saddle point with low training error and good properties are poorly understood. While desirable properties, such as low generalization error, can often be achieved, others, such as robustness to adversarial attacks, remain elusive. Shedding light to the statistical properties of convergence points can lead to the development of fast algorithms that converge to parameters with good property guarantees. Furthermore, understanding the loss-function landscape for GANs and more general adversarial networks will provide the tools to develop efficient learning algorithms for such systems, as well as understand properties, such as vulnerability to adversarial attacks in applications such as self-driving cars or robots working in close proximity to humans.

The topic

The main objective of this PhD is to develop theoretical models of machine learning which capture the landscape properties of realistic machine learning models, while remaining analytically tractable. Developing and studying these models will draw inspiration from methods of statistical physics, many of which have already proven useful for studying machine learning. The work to be carried out will involve 3 main tasks: (i) using tools from statistical physics to study spin-glass models whose energy shares important properties of the loss function for traditional as well as adversarial learning networks; (ii) modeling the training dynamics of one or two neurons as dynamics for sampling statistical physics ensembles and drawing conclusions for the single-neuron experience of training and the induced two-neuron correlations; (iii) building on understanding obtained by the previous tasks to propose modified loss functions and fast algorithms for convergence to saddle points with specific guaranteed propertied.

Background required

Statistical Physics, Random Matrix Theory

Desired qualifications

Information Theory, Mathematical Physics

Supervisors

Primary Supervisor: Aris Moustakas (National and Kapodistrian University of Athens, GR), http://scholar.uoa.gr/arislm

Title **Traffic Networks with learning**

The context

Why it is important: Traffic congestion has tremendous economic and social consequences. Data transport company INRIX estimated the total cost of lost productivity caused by congestion in the US to be 87 billion dollars in 2018. A well-recognized cause of congestion is the drivers' strategic considerations which will soon become more prevalent and transparent with the expansion of self-driving cars. The theory of price of anarchy made huge progress to explain and quantify these phenomena, but this perspective treats traffic systems as fixed entities. However, these systems are the results of algorithms and distributed protocols, and we need not only to understand their behavior, but additionally to come up with design techniques and mechanisms that have improved qualities.

The topic

The main objective of this PhD is to develop mechanisms that produce outcomes of good quality, for routing settings equipped with Machine Learning tools, such as predictions and online learning. The work to be carried out will involve 4 main questions/topics: (1) dynamic settings where the demand rate (number of vehicles) changes with time, (2) settings where agents are online learners, and choose their paths using some regret minimization algorithm, (3) multiple objectives settings in two distinct directions: (a) agents have diverse preferences (e.g., to minimise their travel time or to avoid traffic lights), and (b) the outcome simultaneously satisfies multiple guarantees (e.g., to minimize the social cost and to guarantee low emissions) and (4) cost-sharing settings, where the demand rate dynamically changes.

Background required

Algorithms and Complexity

Desired qualifications

Algorithmic Game theory, Algorithmic Graph Theory, Online Algorithms, Machine Learning

Supervisors
Primary Supervisor: Alkmini Sgouritsa (Athens University of Economics and Business, GR), https://sites.google.com/site/alkminisgouritsa

Field Causality and Fairness in Data Science and Machine Learning

Title **Causal representation learning**

The context

Why is it important: We all hear the moto, correlation does not imply causation. Several examples exist in real life. A Lamborghini being pink is not improbable. In fact, searching for "pink car" on Google image search returns a Lamborghini in the first 10 results (the image shown on the right). This is an example of a simple data correlation. Not all pink cars are Lamborghini's. The real causal relationship is that a car brand determines a car shape. Obviously, we are ignoring other factors that generate the car picture, such as illumination, the background, the position of the car, etc. Getting such complete causal diagram is very hard but we can simplify by considering that illumination/car position is independent of the car brand. Another level of complexity arises in describing car shape. In fact, car shape is a concept composed of hierarchies of other low-level features (e.g. shape of nose, angle of doors, shape of windows, etc). One can think of similar examples in healthcare clinical decision making. Clinicians for example consider a patient's past, laboratory exams, and clinical images in making decisions and predictions about the future. Years of experience have created a mental causal diagram of how these sources of information come together to make a prediction. We want to create models that can understand the causal relationships between different sources of information and help us build better decision tools.

The topic

The main objective of this PhD is to develop causal machine learning models that are multimodal and can combine a variety of information sources. One major aspect is to learn causal representations from unstructured data (e.g. images/video). For example, given images/videos of plates on a table, can a model learn an association that if the table is not there, the plates should be on the floor. But having an AI that can just learn the concept of the table or the plates alone (like the car example above) is already challenging. Integrating graphs in representation and causal representations is a very rapidly growing field, as it can help the AI find mechanisms that explain such associations and hopefully learn such concepts. Another angle is to consider how representation learning extends when additional sources of information are available e.g. what would happen if we had text about the image, or some information telling us that removing the table made other objects (not plates) end on the floor. Of course, we would like to extend these questions into real application settings (e.g. healthcare) but even if we make progress in other (simpler) datasets where we can verify/validate what we learn it would make a considerable contribution to the field.

Background required

Machine learning principles, Bayesian networks, Mathematics and optimisation, Generative modelling

Desired qualifications

Python programming, Familiarity with PyTorch/TensorFlow, Ability to rapidly prototype, Excellent verbal and written communication skills

Supervisors

Primary Supervisor: Sotirios Tsaftaris (The University of Edinburgh, UK), https://vios.science

Title **Generative models for needles in haystacks**

The context

Why it is important: Generative models aim to identify mechanisms that learn given data how to generate new data. They are not new at all: many of us we have heard of Gaussian Mixture Models. Generative models are extremely useful for example to understand mechanisms of how data are generated, to learn representations for data and be useful to downstream tasks, etc. In fact, today we see their use constantly in a variety of fields. New types of models have surpassed performance of previous models. For example while before VAEs and GANs where considered kings, nowadays diffusion models seem to reign supreme. They are powering for example engines such as Stable Diffusion (Stability.AI), ImageGen (Google), and DALL-E2 (OpenAI) where multimodal text-to-image generation is possible. These models have also been expanded to video, 3D and even other exotic type of datasets (e.g. proteomics, molecular architectures, etc). Despite years of advances in generative models, several major questions remain unexplored. How much data do they need to be trained? How do such generative models behave when data are generated in a skewed –non independent-- fashion? Do they memorise rarely occurring patterns or can generalize even from few examples? Answering these questions are imperative to understand the broad applicability of a generative approach to AI-driven decision making.

The topic
The main objective of this PhD is to develop a better understanding of generative models and how they learn from imbalanced data. The work to be carried out will aim to address the questions outlined above. We envision several ways to get there. For example, we might start by identifying suitable benchmark datasets that we know ground truth generating factors and modifying their distributions to see how different generative models behave. Thus, we would explore and train several types of models. Of critical importance will be to identify mechanisms that permit us to evaluate qualitatively and quantitatively these models and. This will then inform directly how different models behave and ergo inform the design of new models or new processes for training generative models in the presence of imbalanced datasets.

Background required
Machine learning principles, Bayesian statistics, Mathematics and optimisation, Generative modeling

Desired qualifications
Python programming, familiarity with PyTorch/TensorFlow, Ability to rapidly prototype, Excellent verbal and written communication skills

Supervisors
Primary Supervisor: Sotirios Tsaftaris (The University of Edinburgh, UK), https://vios.science

Title **Explainable Fairness in Networks**

The context
Why it is important: The AI revolution has resulted in several automated ML-powered systems that make decisions for various aspects of our lives. The increased complexity of these systems has often render them into black boxes. This has led to the emergence of Explainable AI (XAI). This thesis will focus on explainability for algorithmic unfairness, i.e., the case where the output of an ML system is found to be discriminative against an individual, or a group of individuals. Several such cases have been reported recently including unfair behavior of ML systems used in making decisions in the health and justice domain. We will exploit XAI methods to explain the causes of unfairness. Although XAI has received much current attention, explainability of unfairness is quite novel. The focus will be on ML models and algorithms that use graph data. Graphs are a ubiquitous data model for entities, their relations, dependencies, and interactions. There are several real-world graphs, representing networks ranging from social, communication and transportation networks to biological networks and the brain.

The topic
The main objective of this PhD is to develop models and techniques for explaining unfairness of AI algorithms with a focus on graph data. The work will involve the following 3 steps: (1) We will first build on counterfactual explanations, a specific type of explanations that consider the minimal set of changes in the input data that would reverse the decision of the algorithm. We will consider changes in the input that would improve the fairness of a decision. Besides fairness towards individuals, we will also investigate fairness towards groups of individuals. This introduces novel problems, since most work on XAI aims at explaining individual decisions, ignoring the fact that AI decisions can have a collective (side) effect when looked in aggregation. (2) We will extend our explanations for the case of algorithms that operate on graphs, such as graph neural networks. This step introduces several challenges given the complexity of graph data. (3) We will consider additional explanation types, such as contrastive explanations, that compare the output of specific pairs of input data.

Background required
MSc (or, integrated MSc) in computer science, or related field

Desired qualifications
Background in machine learning, graph theory, or data analysis.

Supervisors
Primary Supervisor: Evaggelia Pitoura (University of Ioannina, GR), http://www.cs.uoi.gr/~pitoura

Title **Fairness of Network Algorithms and Processes**

The context
Why it is important: The AI revolution has resulted in several automated ML-powered systems that make decisions for various aspects of our lives, such as the products we buy, the information we consume, the career we follow, or the health care we receive. Given the stronghold these algorithms have on our lives, there are

concerns about possible biases incorporated in the algorithms or the data used for their training. There is a strong interest in ensuring that the operation of the AI algorithms is fair and just. Most of the work on fair AI has focused on classification. Recently there are efforts in extending the work on fairness to other domains, beyond classification. One such domain is network analysis, where the aim is to make network algorithms and processes fair. This is an exciting new field with many open research questions that could also have an impact on society. For example, in a social network, can we increase the influence of minorities by appropriate link recommendations, or by making content distribution fairer? The thesis will focus on such questions.

The topic

The main objective of this PhD is to study fairness in graphs and graph processes and algorithms. The work will involve the following topics: (1) Consider different notions of fairness in networks, such as centrality fairness or community fairness. We will provide models and definitions for fairness and bias in these cases. (2) Understand the temporal evolution of fairness for the different metrics we consider. This involves empirical measurements, but also models for network growth that incorporate bias and unfairness. (3) Mitigate unfairness. This will be approached either by modifying existing algorithms (e.g., centrality computation) to take into account fairness, or by intervening in the network structure through link recommendations to make the underlying network fairer.

Background required

Network Analysis, Graph algorithms and processing, Machine Learning

Desired qualifications

M.Sc. (or integrated M.Sc.) in Computer Science or related field, Python Programming, Linear Algebra

Supervisors

Primary Supervisor: Panayiotis Tsaparas (University of Ioannina, GR), http://www.cs.uoi.gr/~tsap

Title **Representation bias in networks**

The context

Why it is important: The AI revolution has resulted in several automated ML-powered systems that make decisions for various aspects of our lives. Given the stronghold these algorithms have on our lives, there are concerns about the fairness of the decisions that they make. Representation bias is an important source of unfairness. With representation bias, we refer to cases where the input (training) data population under-represents and potentially fails to generalize well for some parts of the target population. Such biases can appear for example during the collection, integration, or sampling of the input data. Although there has been work on tabular data, representation bias for graphs is less understood. Graphs are a ubiquitous data model for entities, their relations, dependencies, and interactions. There are several real-world graphs, representing networks ranging from social, communication and transportation networks to biological networks and the brain.

The topic

The main objective of this PhD thesis is to model, understand the causes and mitigate representation bias in graph data. Handling representation biases in graphs poses novel challenges given the correlation of attribute values with structure, the multiple, potentially continuous attributes, and the case of heterogenous graphs with different types of edges and nodes. The work will cover the following topics: (1) Conceptualize and provide concrete formulations of the several forms of graph representation bias; (2) Propose new algorithms for mitigating graph representation bias for several tasks including data collection from multiple sources, integration, and sampling; (3) Measure and address the effect of representation bias in various machine learning models (e.g., graph neural networks) and for various downstream tasks (e.g., link prediction).

Background required

MSc (or, integrated MSc) in computer science, or related field.

Desired qualifications

Background in machine learning, graph theory, or data analysis.

Supervisors

Primary Supervisor: Evaggelia Pitoura (University of Ioannina, GR), http://www.cs.uoi.gr/~pitoura

Field Machine Learning and Computer Vision

Title **Continual/lifelong visual learning**

The context
Current top-performing deep learning models for visual recognition mostly operate under the ideal conditions of a stationary environment and thus often make totally unrealistic assumptions: for instance, they assume that training data are independent and identically distributed, training samples have been properly selected so as to be equally balanced among all different classes, datasets remain static across time, no distribution shifts are assumed in the incoming input data, and the set of visual classes or tasks to be learnt are predetermined in advance and fixed. This is, of course, in stark contrast with how humans operate who are able to learn in an online and continuous manner in an ever-changing and dynamic environment. It thus comes as no surprise that even the most powerful deep learning models today can often exhibit a significant drop in performance (or even fail completely) when they are confronted with learning scenarios that are dynamic and incremental in nature. Addressing or reducing this fundamental limitation of existing deep learning models would significantly enlarge the domain of applicability, the quality, as well as the robustness of these models, and would help in further increasing their impact even in very challenging areas such as self-driving cars, virtual assistants and robotic automation.

The topic
The main goal of this PhD is to develop novel deep learning-based methods for continual/incremental visual learning, where incoming training data as well as visual tasks are presented to us in a sequential manner over time. The work that will be carried out will explore/revolve around the following questions/topics: (i) How can we develop neural network-based approaches that can successfully deal with catastrophic forgetting, which remains a major challenge and involves the fact that learning of new tasks often comes at the cost of a significant deterioration of neural network performance with respect to older tasks? (ii) How can we design neural network architectures that encode proper inductive biases for continual/incremental learning? (iii) In addition, the goal will be to explore how we can move beyond simple settings and consider more general scenarios for continual learning that can involve, e.g., open-ended recognition, no task boundaries during training, memory budget constraints, as well as having input and output distributions that are distinct or gradually changing between tasks.

Background required
Computer vision/image analysis, deep learning, machine learning

Desired qualifications
Good prior experience with Pytorch or Tensorflow and Python programming

Supervisors
Primary Supervisor: Nikos Komodakis (University of Crete, GR), http://www.csd.uoc.gr/~komod

Title **Large scale unsupervised learning**

The context
Computer vision models based on deep neural networks have shown to provide amazing levels of performance over the last years. One scenario in which they have proved to be particularly successful is for problems where there is an abundance of labeled image data that are of high-quality and carefully curated. However, when such large training datasets are not available or are difficult to collect, they seem to struggle. On the other hand, for many vision tasks it is very often the case that raw, non-annotated data can be readily obtained in large quantities. Consider, for instance, the vast amount of data in the form of images, text or video that is available online. Despite being non-annotated, such large scale multi-modal data contain a tremendous amount of useful information about the structure of visual data and their semantics. If deep neural networks were able to more effectively exploit this type of large scale unlabeled data, this could potentially significantly improve the quality of visual data representations that are extracted by these models and would allow these representations to generalize to a much wider variety of tasks.

The topic
The main objective of this PhD is to circumvent the dependence of deep learning models on large-size manually labeled datasets by allowing them to properly take advantage of large amounts of unlabeled data that are often readily available. To that end, the work that will be carried out will explore a variety of different methodological approaches, covering different aspects of the problem. This includes the development of novel approaches for large scale self-supervised and

weakly-supervised learning for deep neural networks, unsupervised methods that can harness the power and complementarity of multiple modalities (including, e.g., - besides images - video, 3D geometry and text), as well as algorithms that can be used for extracting (in an unsupervised manner) disentangled representations that separate the distinct, informative factors of variations in the data (for single-modal and multi-modal cases).

Background required
Computer vision/image analysis, deep learning, machine learning

Desired qualifications
Good prior experience with Pytorch or Tensorflow and Python programming

Supervisors
Primary Supervisor: Nikos Komodakis (University of Crete, GR), http://www.csd.uoc.gr/~komod

Title **Neuromorphic Algorithms for Visual Motion Perception**

The context
Today's AI paradigm is characterized by high performance using a lot of data and computational power. Biological brains, on the other hand, are characterized by extraordinary efficiency both in power consumption and sample complexity. Recently, a new paradigm of sensing and computation emerged that goes beyond classic sensors and the von-Neumann architecture and offers itself to tasks like visual motion that necessitate a low-power, low-latency response. The neuromorphic paradigm is supported by event-based cameras and asynchronous processors and opens new ways of thinking about algorithms for perception and intelligence.

The topic
Visual motion perception is a crucial brain capability in all organisms since it supports navigation and the detection of targets for prey or avoidance. Event-based cameras provide an almost continuous time-sparse space capture of changes in the visual field. Detecting a robot's change direction, identifying independent object motions, and tracking objects are basic vision capabilities that have been only solved with high-power consumption neural networks implemented on synchronous GPU's. The topic of this research is the design of algorithms and the underlying mathematical study of robustness and optimization using asynchronous networks like Spiking Neural Networks. The research will be developed in two steps: deployment of inference with models that were learned using traditional deep learning and the design of optimization for learning of motion filters using spiking neurons.

Background required
Computer vision, deep learning principles, fundamental neuron models

Desired qualifications
Familiarity with deep learning platforms

Supervisors
Primary Supervisor: Kostas Daniilidis (University of Pennsylvania, USA), https://www.cis.upenn.edu/~kostas/

Title **Symmetry in neuromorphic algorithms and architectures**

The context
AI has achieved superhuman performance in several machine learning tasks, including image recognition and natural language processing and generation. Regardless of this success, the sample complexity and the energy consumption of current AI systems are orders of magnitude higher than the biological brain. There is a need to re-think intelligence across the entire design stack spanning low-power sensing modalities, learning algorithms, and the underlying hardware to produce systems that are efficient in internal data bandwidth and power consumption. Such algorithms can break the power barrier of the current GPU systems and enable embodied edge computing guided by biological principles. While such architectures are efficient in computation, they still need a lot of data to approach the state of the art of deep learning systems. One way to avoid the plethora of data is to avoid data augmentation by building in symmetry in the architectures by applying principles of equivariance and invariance.

The topic
Neuromorphic visual data streams are continuous in time. Binning these event streams into frames defeats the purpose of immediate brain-like processing. Invariance in classical computer vision has been applied only to frames or 3d data. A new mathematical framework is needed that will create representations that are equivariant

to coordinate transformations but also the motion of the agent or scene components. Such mappings have to be made out of infinite impulse response filters in time and be equivariant to spatiotemporal group transformations. Moreover, the sparsity of events lends itself to applying attention mechanisms instead of convolutions that can facilitate the continuous grouping of motions and tracking of objects. Goal of the thesis is to design architectures with neurons that are equivariant to local or global transformations and perform better than approaches that necessitate extensive data augmentation to capture such input variations.

Background required
Machine learning and computer vision principles

Desired qualifications
Familiarity with DL platforms, willingness to study and use differential geometry and harmonic analysis

Supervisors
Primary Supervisor: Kostas Daniilidis (University of Pennsylvania, USA), https://www.cis.upenn.edu/~kostas/

Title **Tensor methods for higher-order deep learning on multimodal/multiway data**

The context
Higher-order neural networks, such as deep polynomial neural networks (DPNs), have been recently proposed as a novel class of deep function approximators based on polynomial expansions. That is, the output is a high-order polynomial of the input. DPNs have demonstrated exceptional performance in discriminative tasks and data generation via adversarial training. The unknown parameters of the polynomial expansion are naturally represented by high-order tensors (i.e., multidimensional matrices). Interestingly, DPNs are quite generic: by applying different tensor decompositions on parameter tensors, we obtain well-known and widely adopted neural network architectures such as non-local neural networks, ResNet, and StyleGAN, among others. Therefore, DPNs act as a unification framework for deep learning architectures and provide a principled way to study and design novel neural networks with desired inductive biases. Also, DPNs, by construction, can capture higher-order multiplicative interactions of all features. Such property has been proved theoretically and empirically as a particularly strong inductive bias when fusing multiple streams of information and when conditional data generation is required. Furthermore, DPNs realize additive polynomial functions, which provide interpretability: lower-degree polynomials that have interactions between a few features are interpretable but likely less accurate, whereas polynomials of higher degree have larger learning capacity at the cost of interpretability.

The topic
We are looking for a passionate and highly motivated PhD student interested in the intersection of tensor methods and deep learning with a focus on higher-order deep learning on multimodal/multiway data. The main objective of this PhD is to develop methods and software tools for higher-order deep neural networks by leveraging multivariate polynomial expansions that capture higher-order auto- and cross-correlations between the input variables from multiple data sources and making such polynomial expansions practical in large-scale data regimes. The work to be carried out will involve 3 main topics: (1) Extensions of linear operations (e.g., convolution operators, attention mechanisms) to appropriate multilinear operations with multiple kernels that naturally apply to higher-order data without flattening (2) Investigate robustness and regularization techniques that preserve important higher-order interactions in higher-order neural networks using randomized tensor decompositions (3) Study the interpretability of additive higher-order deep models. Besides the methodological principles, this PhD project will focus on applications involving large-scale multimodal/multiway data (e.g., computer vision and/or NLP and/or earth observation, etc.).

Background required
Machine learning fundamentals (incl. linear algebra, calculus and probability), Deep learning, Tensor algebra and tensor decompositions

Desired qualifications
University degree in electrical engineering, computer science, or related disciplines, Excellent knowledge of Python and preferably of a deep learning framework (e.g. PyTorch, Jax), Proficiency in English

Supervisors
Primary Supervisor: Yannis Panagakis (National and Kapodistrian University of Athens, GR), http://users.uoa.gr/~yannisp/

Title **Theoretical foundations of geometric multimodal machine learning**

The context
Multimodal machine learning is a multi-disciplinary research field that aims to design intelligent systems that learn to understand, reason, and interact with the natural world through integrating multiple sources of information (e.g., linguistic, acoustic, visual). The application domains are ubiquitous, including human-computer interaction, embodied autonomous agents, media generation (e.g., text-to-image generation), and multisensor fusion in fields such as climate monitoring, neuroscience, pharmacology/drug discovery, bioinformatics, etc. The increasing heterogeneity of data calls for machine learning models that combine multiple inductive biases. However, combining data from various sources is challenging because appropriate inductive bias may vary by data modality. Despite the ongoing success in this field, strikingly, many methods are ad-hoc and rely on the availability of large amounts of annotated data and computational power. However, such requirements are often not met in practice and most of the empirical advancements in this field are observed in data modalities residing on spaces with Euclidean topology (e.g., images), while multimodal learning on arbitrary geometries is still in its infancy.

The topic
We are looking for a passionate and highly motivated PhD student interested in the foundations of multimodal machine learning and geometric deep learning. The PhD project aims to develop a new theory and a unifying computational framework for machine learning on multimodal, or more generally, multi-way data, i.e., data originating from sources of variable heterogeneity and arbitrary geometry/mathematical structure. Key topics that we aim to address are: 1. Universal approximation properties on multiple modalities of arbitrary geometry. How to build architectures invariant/equivariant to intra- and cross-modal symmetries? 2. Theoretical analysis of fusion. Under which conditions multimodal learning improves generalization or signal recovery? What are the advantages and disadvantages of different fusion mechanisms? Does it improve expressive power? 3. Cross-modal interactions. How to formally define cross-modal interactions (potentially of high-order) and introduce appropriate inductive biases in neural network architectures. Besides the methodological principles, a primary application domain of interest is that of generative modeling (and simulation/data-driven control) of multimodal data, with a particular focus on multi-dimensional networks (e.g., in biology, computational neuroscience, and sociology) and manifolds (e.g., for climate modeling).

Background required
Machine learning fundamentals (incl. linear algebra, calculus and probability), Good knowledge of deep learning, Basic knowledge in the following is a plus: graph theory, network science, tensor algebra.

Desired qualifications
University degree in electrical engineering, computer science, or related disciplines, Python programming, deep learning frameworks (e.g. PyTorch, Jax), Proficiency in English

Supervisors
Primary Supervisor: Yannis Panagakis (National and Kapodistrian University of Athens, GR), http://users.uoa.gr/~yannisp/

Title **Visual recognition with limited supervision**

The context
Current state-of-the-art methods for visual recognition still lack several of the key characteristics that represent hallmarks of human intelligence. One of them relates to the remarkable ability of humans to effortlessly learn to recognize novel visual categories based just on a limited set of annotated examples. This is in deep contrast to the amount of supervision that is often required by state-of-the-art models for image recognition. As an example, recent such models based on transformer architectures have to be trained on image datasets such as the JFT300M and JFT-3B, which consist of 300 million images and 3 billion images respectively, in order to be able to reach their state-of-the-art performance. Addressing this fundamental issue of data inefficiency remains an important research problem today, which would allow to dramatically expand the deployment of AI in the real world and would open the way to new applications in areas as diverse and critical as industrial automation and inspection, medical diagnosis and treatment, healthcare and robotics.

The topic
The main objective of this PhD is to develop deep learning methods that will significantly enhance the ability of visual recognition models to generalize from limited supervision. The work that will be carried out will revolve around the following questions/topics: (i) How can we develop more effective neural network-based approaches for few-shot learning, which is a learning paradigm that is based on the concept that reliable algorithms can be created to make predictions from minimalist datasets (e.g., datasets consisting of even just a single training

sample)? (ii) What role do neural network architectures can play regarding generalization efficiency in this case and how can we develop novel neural network architectures adapted specifically for few-shot generalization? (iii) How can we go beyond few-shot learning and also consider other more general forms of limited supervision for visual recognition as well as for more complex visual tasks?

Background required
Computer vision/image analysis, deep learning, machine learning

Desired qualifications
Good prior experience with Pytorch or Tensorflow and Python programming

Supervisors
Primary Supervisor: Nikos Komodakis (University of Crete, GR), http://www.csd.uoc.gr/~komod

Field Machine Learning and Natural Language Processing

Title **Deductive Reasoning in Natural Language**

The context

Traditional deductive reasoning algorithms draw conclusions from a given set of premises (previous knowledge) by iteratively applying inference rules (e.g., modus ponens). They combine a formal meaning representation language (e.g., first-order predicate logic) and a heuristic search algorithm (e.g., A*) to find a sequence of inference steps (proof) that, e.g., gradually add to the given premises new conclusions, until the target conclusion is obtained. These algorithms, however, are often impractical, because of three reasons. (1) Knowledge is available mostly in natural language (NL) (e.g., scientific journals, Wikipedia). To apply symbolic deductive reasoning, this knowledge has to be converted to formal meaning representations inference rules can operate on. Despite progress in semantic parsing, this conversion is a significant source of error. (2) Even perfectly accurate semantic parsing produces different meaning representations for semantically equivalent (or very similar) NL expressions (e.g., synonyms, paraphrases). (3) Numerous premises capturing common-sense knowledge are needed. On another front, Deep Learning (DL) recently led to impressive results, but mostly in perception tasks (e.g., speech recognition, object detection) and tasks (e.g., image/text classification, machine translation) that do not require an explicit multi-step human-interpretable proof of how multiple pieces of information were combined to reach a decision. Related to deductive reasoning is the Natural Language Inference (NLI or Textual Entailment) task. Given a premise P and a hypothesis H, both in NL (e.g., sentences, paragraphs), NLI requires deciding if P implies H (e.g., "Mary is playing in the garden" implies "Mary is outside"), or the negation of H, or none. Compared to symbolic deductive reasoning, NLI has the advantage that P and H are stated in NL. There is no need to convert them to formal meaning representations, and a DL-based NLI system pre-trained on large datasets may also capture common-sense knowledge. However, the NLI task typically corresponds to a single inference step of the proof a deductive reasoning algorithm would search for. Furthermore, both P and H are given, whereas when searching for a proof, intermediate conclusions H need to be generated.

The topic

We aim to combine DL with deductive reasoning from the symbolic Artificial Intelligence tradition, to produce new algorithms capable of reaching conclusions requiring multiple inference steps, using premises and conclusions expressed in NL. Our starting point will be recent work like the FaiRR system (Sanyal et al. 2022), an improvement of ProofWriter (Tafjord et al. 2021). FaiRR assumes a knowledge base (KB) of facts and rules, similar to Horn clauses, but expressed in NL. At each inference step, a DL classifier is fed with a string containing the target conclusion concatenated with all the facts and rules of the KB, and selects the rule to be used. Another DL classifier is given the target conclusion, the selected rule, and all facts, and selects the facts to be combined with the selected rule. A DL encoder-decoder is then given the selected rule and facts, and generates a conclusion in NL (new fact). The inference steps terminate when a generated conclusion matches the target conclusion. FaiRR and similar recent work (Clark et al. 2020, Krishna et al. 2021, Tafjord et al. 2021, Yang et al. 2021, Dalvi et al. 2021, Qu et. al. 2022) illustrate the viability of deductive reasoning in NL, but still have several limitations, including (a) scalability problems with large KBs, (b) suboptimal search algorithms, (c) generating invalid proofs, (d) requiring unrealistic and costly training data (e.g., with gold proof trees), (e) error propagation due to components trained independently, (f) fairness and bias issues (e.g., linking inference steps with gender or ethnicity). The main objective of the two PhD students will be to develop improved NL deductive reasoning algorithms that will address these issues. The new algorithms will be studied in applications like legal judgment prediction (predicting the outcome of a court case), biomedical question answering (QA), and fact verification (fact-checking reported information). Rather than aiming to replace domain experts (e.g., judges, doctors), we will also aim to develop synergistic human-computer deductive reasoning (e.g., proposing relevant facts or next intermediate conclusions to an expert who selects or edits them). The two PhD students will focus on different components and/or approaches of the NL deductive reasoning methods to be developed, and different applications (legal and biomedical), but the two PhD theses will provide feedback to each other.

Background required

BSc and MSc (or integrated MSc) in Computer Science or Computer Engineering. Strong background (completed courses and/or experience) in Artificial Intelligence, Machine Learning, Deep Learning, Natural Language Processing. Proficiency in English.

Desired qualifications

Python programming for NLP and machine learning, with libraries like NLTK, spaCy, scikit-learn. Developing DL models for NLP in PyTorch and/or Tensorflow/Keras. Linux/Unix experience, including shell scripting.

Supervisors

Primary Supervisor: Ion Androutsopoulos (Athens Univ of Economics and Business, GR), https://www.aueb.gr/users/ion/

Title **Greek Computational Dialectology**

The context

Why it is important: No language is a monolith; within the same language variation results from sources such as regional, social class and mode-of-usage differences. However, modern language processing systems tend to ignore such variation and to treat it simply as noise. The result is that systems will make persistent errors for users who speak or write hundreds of other less-common varieties (e.g., Cypriot Greek). The goal of this project is to focus on Modern Greek Dialects and produce corpora and methods needed for adequately handling several Greek varieties with modern neural language models.

The topic

The main objective of this position is to develop annotated corpora for Greek dialectal varieties, use them for comparative studies of language models and infuse neural models with knowledge from linguistics. The work to be carried out will involve 3 main topics: (1) the creation of annotated corpora for Greek varieties, in particular with computationally-aided approaches; (ii) the study of specific dialectal phenomena (focusing on morphology and/or syntax), also through probing language models (iii) the investigation of ways to represent linguistic knowledge and incorporate it into the training of computational models.

Background required

Computational Linguistics, Corpus Annotation and Construction, Machine learning principles

Desired qualifications

Python programming, General notions of dialectology, Morphology and Syntax, Willingness to work with an interdisciplinary team (e.g., learn basic NLP and ML principles)

Supervisors

Primary Supervisor: Stella Markantonatou (Institute for Language and Speech Processing, "Athena" Research Center, GR), https://www.ilsp.gr/en/members/markantonatou-stella-2/

Title **Understanding the Inner Mechanics of Multilingual Language Models**

The context

Why it is important: Large pre-trained neural language models are revolutionizing the field of natural language processing, becoming the default building block for thousands of downstream user-facing applications. However, training such models requires vast amounts of data which are only available for a few dominant languages and language varieties. As a result, systems trained this way will make persistent errors for users who speak or write hundreds of other less-common varieties (e.g. Cypriot Greek). The goal of this project is to stop treating language variation as noise and start modeling it, because millions of diverse users can potentially benefit from natural language technologies, but not if these technologies are only built for standard language varieties.

The topic

The main objective of this PhD is to develop novel, linguistically-motivated techniques for understanding the inner mechanics of training multilingual models, and for incorporating such linguistic knowledge into large neural models, with an approach combining neural and symbolic methods. The work to be carried out will involve 3 main questions/topics: (1) study the training dynamics between languages in multilingual models and probing for linguistic knowledge; (2) perform comparative studies using data from dialectal varieties (e.g., Standard Modern Greek vs Griko vs Lesbian) with the goal of analyzing fine-grained phenomena that may not be adequately modeled by neural models; (3) investigate ways to improve the robustness of LMs to language variation without simply relying on more data.

Background required

Machine learning principles, Natural Language Processing, Deep Learning

Desired qualifications

Python programming, Computational Linguistics, Willingness to work with an interdisciplinary team (e.g., learn basic principles of linguistics and more specifically of dialectology)

Supervisors

Primary Supervisor: Antonios Anastasopoulos (George Mason University, USA), https://cs.gmu.edu/~antonis/

Field Machine Learning and Life Sciences

Title **Causal machine learning for rigorous single- and multi-modal healthcare data analyses**

The context

Why it is important: Multi-modal (i.e.-source) data such as medical imaging (MRI, CT, PET/CT, Ultrasound, x-Rays), molecular profiling (genomic, proteomic, transcriptomic) and patient level data (clinical history, sociodemographics), are becoming substantially important in improving clinical diagnosis and therapy, and in informing decisions and predictions in healthcare (e.g. precision medicine). Although AI is commonly assumed to improve these biomedical procedures, relies heavily on supervision ("labels" or "annotations") and correlations (not causally related outcomes) even when dealing with single-modal data. Moreover, AI models can only understand data on which they have been trained, which compromises unseen and out-of-distribution data analyses. To improve these processes, AI models need to understand which subset of the data forms the cause of their outcomes. Moreover, to be able to combine information from different modalities in AI models, we need to process data jointly and identify these subsets simultaneously across and within modalities. This can lead to more clever AI models that will rely less on explicit supervision and will be able to generalize on unseen single- and multi-modal data. This can revolutionize the way we develop state-of-the-art AI methods in modern healthcare.

The topic

This PhD will aim to develop ML methods that can endow a causal interpretation in predictions in healthcare data. The objectives of this work will aim to address the open questions described above. We aim to explore how we can use causal methods to better understand and describe the data collection process, the feature selection process and ultimately the actual task of decision and prediction. While we will start with structured data (e.g. tabular) one can see interesting complexities that arise in data that are unstructured (e.g. medical images). In that setting, extracting relevant causal representations is important. Thus, we envision this work on real healthcare data to touch on (1) endowing causal understanding to classical ML tasks and data collection for improving robustness; (2) devising and comparing models for causal feature representation and feature selection by analyzing unstructured data; and (3) extending from the single modal to the multimodal setting causal feature selection and inference by co-analyzing unstructured along with structured data.

Background required

Machine learning principles, Bayesian statistics, Mathematics and optimisation, Generative modeling

Desired qualifications

Python programming, familiarity with PyTorch/TensorFlow, Ability to rapidly prototype, Excellent verbal and written communication skills

Supervisors

Primary Supervisor: Giorgos Papanastasiou (Univ of Essex, UK), https://www.essex.ac.uk/people/papan14104/georgios-papanastasiou

Title **Out of distribution detection and domain adaptation of models in the medical domain**

The context

AI-based models are gaining a lot of attention currently since they provide state-of-the-art performance for a variety of tasks and applications. However, these models could be prone to biases and poor generalizability, as it has been demonstrated recently in several studies. Such limitations are crucial, especially for medical applications on which the decisions of such models will affect the life of patients and reduce their clinical use for precision medicine. The problem of poor generalizability is quite important for computational pathology applications since histopathology slides vary a lot across different medical institutions resulting in significant inter-hospital variability at the whole slide level. This indicates the need for the development of methods that generalize well on out-of-distribution samples. Domain adaptation methods could help in this direction and this project aims to develop novel methods focusing on the detection of out-of-distribution samples and bias identification and their robust integration into trained deep-learning models. Multicentric public and private data from Gustave Roussy Hospital in Paris and Henry Dunant Hospital Center in Athens will be investigated to build models for personalized treatment.

The topic

The scope of this project is to enhance the generalizability of the provided models and identify potential biases in the processing of whole slide images that may be due to the properties of the datasets. Specifically, we aim to investigate domain adaptation approaches to enhance the generalizability of models trained on multicentric datasets with highly imbalanced cohorts with respect to the clinical endpoint at hand. In particular, methods similar to the multi-domain image-to-image translation models will be explored for training and test data augmentation together with adversarial attacks to generate guided perturbations for the generated domains, strengthening the generalization of the models. Moreover, possible biases concerning the center at which the data were collected, the sex, and other clinico-biological attributes will be identified, and models will be extended to provide bias correction alternatives. Since our goal is to use our models in clinical practice, this is a

Background required

MSc or BSc degree in engineering, science, mathematics, or computer science, Expertise in programming and Python, Clear interest in artificial intelligence and medical image analysis

Desired qualifications

Medical Data Knowledge, Deep Learning Frameworks

Supervisors

Primary Supervisor: Maria Vakalopoulou (CentraleSupelec, FR), https://scholar.google.gr/citations?user=FKUHYqMAAAAJ&hl=en

Title **Self- and cross-attention models for biomarker discovery on multimodal data for cancer diagnosis**

The context

Histopathology tissue slides are predominantly used for tumor phenotyping and hence patient stratification and treatment adoption. These slides contain biopsies of pathological tissue from patients and are typically stained to highlight different tissue structures or cancer cell characteristics. Experienced pathologists assess such tissue slides with the use of a microscope aiming to quantify the different biomarkers. Even though this process allows for evidence-based treatment selection, it is prone to inter-reader variability, and it does not provide usable insights on correlations between patient-specific data and treatment outcomes. Artificial intelligence has been shown to address such issues; however, there are still a lot of challenges. In particular, current algorithms have been handicapped to replicating what humans can do instead of trying to go beyond that, mainly due to methodological limitations and knowledge gaps around AI and high dimensional data. To this end, this research project aims to design and evaluate novel multimodal learning schemes that incorporate spatial dependencies on the digitized tissue slides toward predicting treatment response outcomes. Our goal is to propose novel attention mechanisms for multimodal fusion between tissue biopsies, and genetic and clinico-biological data towards a better understanding of cancer. Public and private data from Gustave Roussy Hospital in Paris and Henry Dunant Hospital Center in Athens will be investigated to build models for personalized treatment.

The topic

In this project, we will investigate transformer-based architectures that utilize self-attention modules towards biomarker discovery. Such models can provide task-specific saliency maps concerning different endpoints, which can serve as computational staining schemes and can assist in the identification of novel biomarkers. To stabilize training due to the small amount of available data and enhance performance, we will explore self-supervised pre-training frameworks with self-distillation. Specifically, we aim to investigate transformer-based cross-modality attention models for gigapixel images and genetic data toward biomarker discovery. Lastly, recent multimodal schemes based on transformers and generative models will be investigated and extended into novel schemes for biomarker discovery.

Background required

MSc or BSc degree in engineering, science, mathematics, or computer science, Expertise in programming and Python, Clear interest in artificial intelligence and medical image analysis

Desired qualifications

Medical Data Knowledge, Deep Learning Frameworks

Supervisors

Primary Supervisor: Maria Vakalopoulou (CentraleSupelec, FR), https://scholar.google.gr/citations?user=FKUHYqMAAAAJ&hl=en

Title **Unsupervised quantification and biomarker prediction in computational pathology for limited amount of data**

The context

Cancer is a leading cause of death worldwide, accounting for nearly 10 million deaths in 2020 or nearly one in six deaths overall. Microscopic assessment of cancer tissue by pathologists is the golden standard of diagnostic and treatment decision-making for patients. However, the current assessment protocols, referred to as cancer grading systems, are limited to features that can feasibly be extracted by humans, leaving vast amounts of information captured in cancer tissue unused. To this end, machine learning and deep learning algorithms have already shown that it is capable of improving the diagnostic performance and efficiency of pathologists in the automatic quantification of different biomarkers which mainly describe the tumor microenvironment and the cancer subtype (e.g., percentage of tumor-infiltrating lymphocytes or PD-L1 expression levels), however, there are still a lot of challenges. These challenges focus on the limited annotations usually available on such applications. In this project, innovative learning algorithms and the latest advances in self/weak supervision, vision transformers, and generative models will be developed to make deep neural networks more effective in medical applications and endow them with transparency and explainability. Public and private data from Gustave Roussy Hospital in Paris and Henry Dunant Hospital Center in Athens will be investigated to build models for personalized treatment.

The topic

Within the scope of this project, tailored deep learning architectures will be developed focusing on spatially aware models for gigapixel-size histopathological whole slide images. Typically whole slide histopathology imaging are of huge size, making the application of machine and deep learning algorithms very challenging. At the same time, the limited amount of annotations that are usually available on such datasets makes the training of such models not trivial. On these grounds, appropriate learning schemes that utilize generative and contrastive schemes will be investigated and coupled with attention mechanisms for biomarker quantification extending classical multi-instance learning schemes. Finally, greedy-based methods for training multiple gradient-isolated modules will be developed towards models for whole slide-level biomarker discovery.

Background required

MSc or BSc degree in engineering, science, mathematics, or computer science, Expertise in programming and Python, Clear interest in artificial intelligence and medical image analysis

Desired qualifications

Medical Data Knowledge, Deep Learning Frameworks

Supervisors

Primary Supervisor: Maria Vakalopoulou (CentraleSupelec, FR), https://scholar.google.gr/citations?user=FKUHYqMAAAAJ&hl=en

Title **Artificial Intelligence (AI) through Network Systems Neuroscience I**

The context

Artificial Intelligence (AI) has been guided by neuroscience: the stochasticity in the neuronal firing has shaped the development of successful regularization that support generalization beyond training data (e.g., dropout), reinforcement learning has been inspired by research in animal learning, and selective visual attentional mechanisms that benefit behavior have influenced AI architectures. Researchers of spiking neural networks (SNNs) have also employed aspects exhibited in real neural circuits of the brain, including analog computation, low power consumption, fast inference, event-driven processing, online learning, and massive parallelism, to build neuromorphic software and hardware platforms. The majority of SNNs have been limited to simple and shallow architectures. The notion of time differentiates ANN and SNN operation: while ANN inputs are static, SNNs operate based on dynamic binary spiking inputs as a function of time. The brain is much more complex than current AI algorithms: e.g., biological neurons are more than non-linear functions. They spike, are leaky, can be stochastic, oscillate and synchronize with different functional modules, and integrate signals from different areas. Similarly, biological synapses are more than analog weights (e.g., with leaky memories, stochastic, operating in different time scales). Biological principles can bring additional features to artificial neural networks and enable the development of more sophisticated simulation platforms that can further guide neuroscientists in new experiments, forming a **multi-disciplinary, cross-fertilization loop**.

This research is driven by the desire to decipher how the brain performs the complicated computations that allow us to learn about the environment and how the interlaced activity patterns across the neocortical populations of neurons bestow the capacity to interact intelligently with the environment and make robust inferences about the world from limited data and in noisy environments.

The topic

The main objective of this Ph.D. is to develop and analyze AI architectures based on biological principles. Specifically, the Ph.D. student will design and implement AI architectures to model biological neural network input and output layers, train them using the extensive observed datasets under various conditions, and examine their performance by employing metrics, such as robustness under noise, classification accuracy prediction, joint information transfer and energy efficiency, and training size. The research will consider requirements about invariance (e.g., in size, and rotation). One of the focuses will be the integration of connectivity patterns identified in the biological networks on the AI architectures and the analysis of their impact on the performance. The proposed architectures will be comparatively analyzed with pre-trained AI architectures with state-of-the-art performance in the domain of vision.

Background required

Machine learning; Network analysis; No biological expertise is required

Desired qualifications

Undergraduate studies in CS, EE, or Mathematics; M.Sc. in data science or computational neuroscience or CS; Strong data science, graph theory, and linear algebra background; Strong programming experience in python/matlab/R; Familiarity with deep learning platforms/frameworks;

Supervisors

Primary Supervisor: Maria Papadopouli (University of Crete, GR) http://www.csd.uoc.gr/~maria

Title **Artificial Intelligence (AI) through Network Systems Neuroscience II**

The context

Artificial Intelligence (AI) has been guided by neuroscience: the stochasticity in the neuronal firing has shaped the development of successful regularization that support generalization beyond training data (e.g., dropout), reinforcement learning has been inspired by research in animal learning, and selective visual attentional mechanisms that benefit behavior have influenced AI architectures. Researchers of spiking neural networks (SNNs) have also employed aspects exhibited in real neural circuits of the brain, including analog computation, low power consumption, fast inference, event-driven processing, online learning, and massive parallelism, to build neuromorphic software and hardware platforms. The majority of SNNs have been limited to simple and shallow architectures. The notion of time differentiates ANN and SNN operation: while ANN inputs are static, SNNs operate based on dynamic binary spiking inputs as a function of time. The brain is much more complex than current AI algorithms: e.g., biological neurons are more than non-linear functions. They spike, are leaky, can be stochastic, oscillate and synchronize with different functional modules, and integrate signals from different areas. Similarly, biological synapses are more than analog weights (e.g., with leaky memories, stochastic, operating in different time scales). Biological principles can bring additional features to artificial neural networks and enable the development of more sophisticated simulation platforms that can further guide neuroscientists in new experiments, forming a multi-disciplinary, cross-fertilization loop. This research is driven by the desire to decipher how the brain performs the complicated computations that allow us to learn about the environment and how the interlaced activity patterns across the neocortical populations of neurons bestow the capacity to interact intelligently with the environment and make robust inferences about the world from limited data and in noisy environments.

The topic

The main objective of this Ph.D. is to develop variants of SNN with different architectures in which the connectivity between their nodes is guided by our findings in the functional networks observed in the visual cortex under different conditions. The SNNs will incorporate also different types of nodes. The Ph.D. student will examine their performance by employing metrics, such as robustness under noise, classification accuracy prediction, joint information transfer and energy efficiency, and training size. The research will consider requirements about invariance (e.g., in size, and rotation). The proposed architectures will be comparatively analyzed with pre-trained AI architectures with state-of-the-art performance in the domain of vision.

Background required

Machine learning; Network analysis; No biological expertise is required

Desired qualifications

Undergraduate studies in CS, EE, or Mathematics; M.Sc. in data science/computational neuroscience/EE/CS; Strong data science, graph theory, linear algebra background; strong programming experience in python/matlab/R; familiarity with deep learning platforms/frameworks;

Supervisors

Primary Supervisor: Maria Papadopouli (University of Crete, GR) http://www.csd.uoc.gr/~maria