My primary research interests focus on **Machine Learning and Optimization**, and their applications in **Energy Systems**. Being a graduate from **Indian Institute of Technology (IIT), Dhanbad** with majors in **Electrical Engineering**, I have always been driven towards working on innovations related to Smart-Grids. The documentary on AlphaGo and its portrayal of how artificial intelligence (AI) can delve into the uncertain corners beyond what human minds can think of, hooked me to the fascinating **applications of reinforcement learning (RL)** in the energy domain. More recently, my research vision got shaped by the experience at **Microsoft Research (MSR), India** where my work revolves around building **"Technology for Emerging Systems"**. Although RL has several shortcomings in its real-world deployability, I believe working on defining safe policies and constrained action value pairs can potentially solve several complex problems in electrical energy sciences. Hence, through graduate studies within the network and collaborations with researchers across Europe, I plan to develop reliable and affordable AI-enabled solutions, robust-enough to drive sustainability goals in the next-generation energy system dynamics.

In the recent past, with the evolution of artificial intelligence (AI), there has been a lot of interests on how to cater to the needs of modern Power System. At MSR, I work with Dr. Akshay Nambi, Tanuja Ganu and Dr. S. Kalyanaraman on applying machine learning for optimization and control in distributed energy resources. Specifically, my work explores various RL techniques, stochastic and traditional optimization algorithms that solves the physics based constraints associated with power systems. Initially, I worked on a **carbon arbitrage (CA) scenario** involving reduction of carbon footprints of the customers of a **partner organization of Microsoft in Ireland**, where scheduling the batteries to charge and discharge accurately at correct time periods plays a vital role. Here, I developed a **gym-compatible environment** that captures the dynamics of the battery, including the battery degradation process, updating the state of charge of the battery and other configurations. Then, I compared the yearly carbon savings across various algorithms: Dual Annealing (DA), Mixed Integer Linear Programming (MILP) and Model Predictive Control (MPC), RL algorithms such as Deep Q-network (DQN) and its variants such as duelling double DQN, and proximal policy optimization (PPO), and a simple baseline: a percentile-based Heuristics model. These specific RL algorithms accurately suit the problem because of the discrete action space variable required by the battery i.e. to charge, to discharge, or to stay idle. The results show that the neural network based policy learners outperform the traditional algorithms in presence of forecast errors with sufficient training data.

Energy operators face several challenges while applying classical RL to large-scale real-world environments because of its learning-by-doing nature. The amount of data collected online while interacting with the environment may also be limited, rising the difficulties of RL agent to learn anything useful. Therefore, I worked on **OfflineRL** or **BatchRL**, namely on **Conservative Q-learning (CQL)** algorithm to solve the problem for the CA scenario. I proposed to first use MPC to solve the objective functions of the scenario by maintaining the constraints on the historically available datasets and then make our own **Markov Decision Process(MDP) based Datasets** for the data-driven offlineRL approach which is used like a supervised learning dataset. The performance boost of OfflineRL results as compared to a simple RL agent trained from scratch showed the superiority of the procedure. This led to a **successful integration** of an example scenario to our Decision Management framework, **"EnCortex"**. More recently, my focus has been on developing the **core energy abstractions**, paired with **modular and extensible optimization algorithms** for **"EnCortex"**. Also, my current vocation involves making the framework compatible enough to handle a range of use cases in the energy domain. It is, now, in the process of being offered to a **large global energy management partner of Microsoft** for wide-scalability and is under patent review rights.

The research community has given a great deal of attention to Learning from Demonstations (LfD) in autonomous agents. Getting a deployed agent to interact with the environment for solving its exploration-exploitation trade-offs can prove to be detrimental in critical systems. Hence, to have the environment interactions performed as a pre-processing step, I experimented on **imitation learning** or more specifically **behaviour cloning**, where I used a trained RL expert to extend its learning to a new agent working on a similar environment. Here, the new agent tries to learn the optimal policies by imitating the expert's decisions i.e. a list of demonstrations sampled from the expert transitions. This in turn helps the energy operators to train the optimizers without having the need of large datasets. The results being promising, with "**EnCortex**" as the main focus of providing a novel one-stop solution for Energy Decision Management Operators, the work has been submitted for a publication at **USENIX NSDI'23**[1] which is under review.

I was driven to pursue research much before joining MSR. In my Sophomore year, I was first introduced to

the exciting line of research in **hybrid distributed generation systems**. More importantly, with the integration of advanced computing and communication technologies, I wished to devise methods on enhancing the efficiency and reliability of future power systems. To further explore my interests,

- I worked with Dr. Swades De as an intern at **IIT Delhi** where, I carried out research in the areas of fault analysis in distribution systems and worked on **Synchro-phasor measurement units (PMU)** based anti-islanding mechanisms within micro-grids. With different types of faults, varying the fault resistances and the location of fault inducement in the buses, I generated a **large fault-detection Dataset** for later research.

- For my Bachelor's thesis at **IIT Dhanbad** under the supervision of Dr. K. Chatterjee and Dr. B. K. Naick, I extended the work done at IIT Delhi, to classify whether islanding has occured or not by accurately shaping the rewards of **Asynchronous Advantage Actor-Critic (A3C)** algorithm and showed a net boost in the overall performance of the method as compared to the state-of-the-art neural network models.

- Continued to being motivated in solving the problem, I worked with Dr. Soham Dutta on hosting capacity (HC) amendment in distribution system to achieve energy sufficiency. Here, firstly for noise cancellation, I use **Spectral-Kurtosis (SK)** to extract the features of the faulted transient signals. Then, based on the optimized feature importance values, I feed in the extracted features to **Histogram-based Gradient Boosting (HGB)** algorithm for accurately predicting the fault type and location with improved confidence. The work has been submitted to **Electric Power Systems Research Journal**[2].

I realize the need for a strong theoretical foundation to pursue advanced research. In this direction, I have always striven for academic excellence - during Bachelors, I graduated as the **Silver Medalist** from my Batch. Working with research professionals has helped me how I can better interpret and address the reviews and comments in paper submissions. I believe that having valuable soft skills like time management and teamwork are vital for survival in challenging academic settings.

I am really excited to contribute to the fields of optimization and control alongwith its applications in energy systems. I would like to grow my knowledge in the algorithmic foundations of convex, non-convex optimization, stochasticity and game theory applications in control and dynamical systems. Over the years I have been influenced by the exciting research done by **Dr. Niao He**. Her recent work on TiAda[3], that achieves near-optimal complexities in deterministic and stochastic settings of non-convex concave minimax problems, interests me a lot. I see potential applications of it in simultaneous decision-making of volume allocation and market-bidding. Having interests in how to design safe RL techniques, I want to delve deeper into the theoretical aspects of RL and have a better understanding of the risk averse, constrained and worst case criteria to be included in the exploration process. I have a strong desire to work with **Dr. David Krueger** in this regard. His outstanding works on reward hacking[4] and objective robustness inspires me a lot to dig deeper into the foundations of RL so as to fruitfully apply it to energy systems. I also aspire to work with **Dr. Volkan Cevher, Dr. Martin Jaggi and Dr. Nicolas Flammarion** in problems involving various aspects of optimization and machine learning. I believe having this theoretical foundation would help me push towards my future goal of building next-generation energy systems to drive sustainability goals.

In summary, I believe I bring with me research experience, industry-tested programming skills, soft skills, and, most importantly, an unquenchable thirst for knowledge and excellence. Therefore, a PhD in Electrical and Computer Engineering from EPFL Graduate School is the next significant achievement I eagerly anticipate in my life.

[1] EnCortex: A General, Extensible and Scalable Framework for Decision Management in New-age Energy Systems, Vaibhav Balloli*, **Millend Roy***, Anupam Sobti, Tanuja Ganu, Akshay Nambi. In: *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI'23) (Under Review)*.

[2] Machine Learning Based Adaptive Fault Diagnosis considering Hosting Capacity Amendment in Active Distribution Network, Sourav Kumar Sahu, **Millend Roy**, Soham Dutta, Debomita Ghosh, Dusmanta Kumar Mahanta. In: *Electric Power Systems Research (EPSR'23) (Under Review)*

[3] TiAda: A Time-scale Adaptive Algorithm for Nonconvex Minimax Optimization, Xiang Li, Junchi Yang, Niao He, arXiv, 2022.

[4] Defining and Characterizing Reward Hacking, Joar Skalse, Nikolaus H. R. Howe, Dmitrii Krasheninnikov, David Krueger, arXiv, 2022