

# About Us

Algorithmic Intelligence Lab (AIL) focuses on the theoretical foundations and practical impact of intelligent systems. Our research centers on *optimization*, *algorithms*, and *mathematical modeling*, and on how these core principles drive advances in *artificial intelligence*, *machine learning*, *social/complex network analysis*, and *data science*.

At AIL, we view intelligence through an algorithmic lens. We develop new algorithmic frameworks, optimization methods, and theoretical guarantees that address fundamental challenges such as scalability, robustness, efficiency, and interpretability. These foundations are tightly coupled with applications in modern data-driven systems, including learning on graphs and networks, large-scale inference, decision-making under uncertainty, and socially informed AI.

A defining feature of AIL is the bidirectional flow between theory and practice. Real-world problems motivate new theoretical questions, while rigorous analysis informs the design of reliable and principled algorithms. Through interdisciplinary collaboration and a strong commitment to mathematical rigor, the Algorithmic Intelligence Lab seeks to contribute foundational insights and practical tools that shape the future of AI and data science.

## 1. Submodular Optimization and Algorithms

*Tag: Theory and algorithms*

Submodular optimization lies at the core of many discrete decision-making problems arising in machine learning, data mining, and network science. Submodular functions capture the principle of diminishing returns and provide a powerful modeling framework for problems involving coverage, diversity, robustness, and information gain. The Algorithmic Intelligence Lab investigates the theoretical foundations and algorithmic design of submodular optimization under various constraints, including cardinality, matroid, knapsack, and more complex combinatorial structures.

Our research focuses on developing efficient and scalable algorithms with provable approximation guarantees, as well as understanding the fundamental limits of optimization in both offline and online settings. We study classical greedy methods alongside modern advances such as streaming algorithms, stochastic optimization, and adaptive and distributed approaches. We also investigate how fairness constraints and group-level objectives can be incorporated into submodular formulations, and how algorithmic guarantees change under such constraints.

Beyond algorithm design, we also analyze complexity, robustness, and stability properties that are critical for deployment in real-world systems, including active learning, feature selection, influence maximization, and network analysis. This line of work establishes a common theoretical backbone that connects the lab's learning, network, and data science research.

## Representative Papers

1. Nemhauser, George L., Laurence A. Wolsey, and Marshall L. Fisher. "An analysis of approximations for maximizing submodular set functions—I." *Mathematical programming* 14, no. 1 (1978): 265-294.
2. Fujishige, Satoru. *Submodular functions and optimization*. Vol. 58. Elsevier, 2005.
3. Krause, Andreas, and Daniel Golovin. "Submodular function maximization." *Tractability* 3, no. 71-104 (2014): 3.
4. Buchbinder, Niv, and Moran Feldman. "Submodular functions maximization problems." In *Handbook of approximation algorithms and metaheuristics*, pp. 753-788. Chapman and Hall/CRC, 2018.
5. Bilmes, Jeff. "Submodularity in machine learning and artificial intelligence." *arXiv preprint arXiv:2202.00132* (2022).
6. Li, Wenxin, Moran Feldman, Ehsan Kazemi, and Amin Karbasi. "Submodular maximization in clean linear time." *Advances in neural information processing systems* 35 (2022): 17473-17487

## 2. Active Learning

*Tag: Learning & Applications of Sec. 1*

Active learning aims to reduce labeling costs by intelligently selecting the most informative data points to query. At the Algorithmic Intelligence Lab, we study active learning through a principled optimization perspective, with a particular emphasis on *submodular objective functions* that quantify information gain, uncertainty reduction, or model improvement.

We formulate active learning as a sequential decision-making problem, where each queried label provides diminishing marginal benefit as more data are observed. This naturally leads to submodular and adaptive submodular formulations, enabling the use of greedy and approximation algorithms with theoretical performance guarantees. Our work explores both pool-based and streaming settings, as well as extensions to structured data, graphs, and large-scale learning systems.

Beyond algorithmic design, we analyze the trade-offs between exploration and exploitation, robustness to noise, and scalability in high-dimensional regimes. By grounding active learning in submodular optimization theory, we aim to develop methods that are not only empirically effective but also mathematically principled. This line of research bridges learning theory, optimization, and real-world applications where labeled data are scarce or expensive.

## Representative Papers

1. Wei, Kai, Rishabh Iyer, and Jeff Bilmes. "Submodularity in data subset selection and active learning." In *International conference on machine learning*, pp. 1954-1963. PMLR, 2015.

2. Kothawade, Suraj, Nathan Beck, Krishnateja Killamsetty, and Rishabh Iyer. "Similar: Submodular information measures based active learning in realistic scenarios." *Advances in Neural Information Processing Systems* 34 (2021): 18685-18697.
3. Golovin, Daniel, and Andreas Krause. "Adaptive submodularity: Theory and applications in active learning and stochastic optimization." *Journal of Artificial Intelligence Research* 42 (2011): 427-486.
4. Akcin, Oguzhan, Orhan Unuvar, Onat Ure, and Sandeep P. Chinchali. "Fleet active learning: A submodular maximization approach." In *7th Annual Conference on Robot Learning*. 2023.
5. Li, Shibo, Jeff M. Phillips, Xin Yu, Robert Kirby, and Shandian Zhe. "Batch multi-fidelity active learning with budget constraints." *Advances in Neural Information Processing Systems* 35 (2022): 995-1007.

### 3. Feature Selection

Tag: *Learning & Applications of Sec. 1*

Feature selection is a fundamental problem in machine learning and data science, with the goal of identifying a small, informative subset of features that preserves predictive power while improving interpretability and efficiency. At AIL, we approach feature selection using *submodular optimization frameworks* that model relevance, redundancy, and diversity in a unified way.

Many commonly used criteria such as mutual information, entropy-based measures, and coverage objectives, exhibit submodularity or approximate submodularity. We leverage this structure to design efficient approximation algorithms with provable guarantees, even in high-dimensional or large-scale settings. Our research also investigates constrained variants of feature selection, including budgeted, structured, and group-based formulations.

In addition to algorithmic development, we study theoretical properties such as approximation bounds, stability, and generalization performance. We also explore connections between feature selection, sparsity, and interpretability in modern machine learning models. By combining rigorous optimization theory with practical learning considerations, this work provides scalable and reliable tools for feature selection across diverse data science applications.

#### Representative Papers

1. Bao, Wei-Xuan, Jun-Yi Hang, and Min-Ling Zhang. "Submodular feature selection for partial label learning." In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 26-34. 2022.
2. Khanna, Rajiv, Ethan Elenberg, Alex Dimakis, Sahand Negahban, and Joydeep Ghosh. "Scalable greedy feature selection via weak submodularity." In *Artificial Intelligence and Statistics*, pp. 1560-1568. PMLR, 2017.

3. Das, Abhimanyu, and David Kempe. "Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection." *arXiv preprint arXiv:1102.3975* (2011).
4. Bateni, MohammadHossein, Lin Chen, Hossein Esfandiari, Thomas Fu, Vahab Mirrokni, and Afshin Rostamizadeh. "Categorical feature compression via submodular optimization." In *International Conference on Machine Learning*, pp. 515-523. PMLR, 2019.
5. Krause, Andreas, H. Brendan McMahan, Carlos Guestrin, and Anupam Gupta. "Robust Submodular Observation Selection." *Journal of Machine Learning Research* 9, no. 12 (2008).
6. Iyer, Rishabh, Ninad Khargonkar, Jeff Bilmes, and Himanshu Asnani. "Generalized submodular information measures: Theoretical properties, examples, optimization algorithms, and applications." *IEEE Transactions on Information Theory* 68, no. 2 (2021): 752-781.

## 4. Recommender Systems

*Tag: Applications of Sec. 1*

Recommender systems sit at the intersection of algorithmic foundations and high-impact applications, making them a natural fit for AIL’s “theory → practice” pipeline. At AIL, we study recommendation through the lens of optimization and learning, emphasizing principled objectives (utility, diversity, exploration) and scalable algorithms (e.g., matrix factorization, neighborhood methods, and modern learning-based recommenders). This thrust cross-connects with feature selection (representation design and sparsity/structure), active learning (efficient preference elicitation and exploration–exploitation), and modeling in complex networks (graph structure, homophily, diffusion-like dynamics that shape exposure and adoption).

We are especially interested in rigorous formulations of constrained recommendation (budgets, diversity, group fairness, and policy constraints) and in understanding how algorithmic choices behave under distribution shift and feedback loops. Many recommender problems also couple naturally with submodular optimization when optimizing coverage/diversity over item sets or user groups, enabling approximation guarantees aligned with our submodular-algorithms core.

### Representative Papers

1. Parambath, Shameem Puthiya, Nishant Vijayakumar, and Sanjay Chawla. "Saga: A submodular greedy algorithm for group recommendation." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1. 2018.
2. Monemizadeh, Morteza. "Dynamic submodular maximization." *Advances in Neural Information Processing Systems* 33 (2020): 9806-9817.
3. Yang, Diyi, Jingbo Shang, and Carolyn Penstein Rosé. "Constrained question recommendation in MOOCs via submodularity." In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pp. 1987-1990. 2014.

4. Chen, Lin, Andreas Krause, and Amin Karbasi. "Interactive submodular bandit." *Advances in Neural Information Processing Systems* 30 (2017).
5. Tschischek, Sebastian, Adish Singla, and Andreas Krause. "Selecting sequences of items via submodular maximization." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1. 2017.
6. Nikolakaki, Sofia Maria, Alina Ene, and Evimaria Terzi. "An efficient framework for balancing submodularity and cost." In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pp. 1256-1266. 2021.

## 5. Data Summarization

Tag: *Applications of Sec. 1*

Data summarization at AIL spans text/document summarization and broader data subset selection for efficient learning and analysis. The unifying idea is to construct compact representations, e.g., summaries, prototypes, or coresets, that preserve the essential signal for downstream tasks. This topic tightly connects to our core research in submodular optimization and algorithms, where diminishing returns provides a principled way to model coverage, diversity, and non-redundancy; it also links to active learning (choosing what to label/read next), feature selection (choosing what to keep), and modeling in complex networks (summarizing graph-structured data and diffusion-relevant content).

On the methods side, we study extractive summarization objectives grounded in submodular functions (coverage/diversity mixtures), scalable maximization under budget constraints, and learning-to-summarize frameworks that tune objective weights to match task metrics. We also emphasize evaluation rigor and robustness: how summaries behave under noise, drift, and domain shift, and how they support reliable decision-making in ML pipelines and network analytics. We also consider representation constraints and balanced-coverage objectives (often compatible with constrained submodular optimization) to ensure summaries are both informative and equitable.

### Representative Papers

1. Mirzasoleiman, Baharan, Ashwinkumar Badanidiyuru, and Amin Karbasi. "Fast constrained submodular maximization: Personalized data summarization." In *International Conference on Machine Learning*, pp. 1358-1367. PMLR, 2016.
2. Badanidiyuru, Ashwinkumar, Baharan Mirzasoleiman, Amin Karbasi, and Andreas Krause. "Streaming submodular maximization: Massive data summarization on the fly." In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 671-680. 2014.

3. Kazemi, Ehsan, Morteza Zadimoghaddam, and Amin Karbasi. "Scalable deletion-robust submodular maximization: Data summarization with privacy and fairness constraints." In *International conference on machine learning*, pp. 2544-2553. PMLR, 2018.
4. Mirzasoleiman, Baharan, Morteza Zadimoghaddam, and Amin Karbasi. "Fast distributed submodular cover: Public-private data summarization." *Advances in Neural Information Processing Systems* 29 (2016).
5. Zhao, Yang, and Tommy WS Chow. "Opinion Summarization via Submodular Information Measures." *IEEE Transactions on Knowledge and Data Engineering* 35, no. 11 (2023): 11708-11721.
6. Yuan, Jing, and Shaojie Tang. "Group fairness in non-monotone submodular maximization." *Journal of Combinatorial Optimization* 45, no. 3 (2023): 88.
7. Li, Wenxin, Moran Feldman, Ehsan Kazemi, and Amin Karbasi. "Submodular maximization in clean linear time." *Advances in neural information processing systems* 35 (2022): 17473-17487.

## 6. Influence Diffusion in Social Networks

*Tag: Social networks analysis & Applications of Sec. 1*

Influence diffusion studies how information, behaviors, or innovations spread through social networks. A central problem in this area is influence maximization: selecting a limited set of seed nodes to maximize the expected spread. At AIL, we study influence diffusion using *submodular optimization methods* grounded in network and algorithmic theory.

Under widely used diffusion models, such as the independent cascade and linear threshold models, the expected influence function is submodular. This property enables efficient greedy algorithms with strong approximation guarantees. Our research extends these classical results by addressing challenges such as scalability to massive networks, uncertainty in network structure, dynamic diffusion processes, and fairness-aware influence strategies.

We also investigate adaptive and online influence maximization, where decisions are made sequentially as diffusion unfolds. Beyond maximization, we study influence control, containment, and competing diffusion processes. By combining submodular optimization theory with realistic network models, this work advances both the theoretical understanding and practical applicability of influence diffusion in social, information, and economic networks.

### Representative Papers

1. Kempe, David, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network." In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 137-146. 2003.

2. Kempe, David, Jon Kleinberg, and Éva Tardos. "Influential nodes in a diffusion model for social networks." In *international colloquium on automata, languages, and programming*, pp. 1127-1138. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.
3. Borgs, Christian, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. "Maximizing social influence in nearly optimal time." In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pp. 946-957. Society for Industrial and Applied Mathematics, 2014.
4. Tang, Youze, Yanchen Shi, and Xiaokui Xiao. "Influence maximization in near-linear time: A martingale approach." In *Proceedings of the 2015 ACM SIGMOD international conference on management of data*, pp. 1539-1554. 2015.
5. Mossel, Elchanan, and Sebastien Roch. "On the submodularity of influence in social networks." In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pp. 128-134. 2007.

## 7. Community Detection in Complex Networks

Tag: *Algorithms & Social/complex networks*

Community detection(clustering) seeks to uncover latent structure in networks by identifying groups of nodes with strong internal connectivity or shared functional roles. At AIL, we study community detection from an algorithmic and optimization-driven perspective, integrating graph theory, statistical modeling, and scalable computation.

Our research explores both classical and modern formulations, including modularity-based methods, spectral algorithms, and probabilistic models. We focus on developing algorithms that are theoretically grounded, computationally efficient, and robust to noise and sparsity. Particular attention is given to large-scale and dynamic networks, where communities evolve over time and exact optimization becomes infeasible.

We also study interpretability, validation, and the relationship between community structure and downstream tasks such as prediction, diffusion, and learning on graphs. In addition, our work explores fairness-aware community detection criteria that promote balanced representation and avoid structural bias. By integrating fairness considerations into network analysis, this research complements the lab's efforts in fair influence diffusion and equitable learning, reinforcing a unified vision of responsible, theory-grounded network science.

### Representative Papers

1. Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." *Physical review E* 69, no. 2 (2004): 026113.
2. Newman, Mark EJ. "Modularity and community structure in networks." *Proceedings of the national academy of sciences* 103, no. 23 (2006): 8577-8582.

3. Von Luxburg, Ulrike. "A tutorial on spectral clustering." *Statistics and computing* 17, no. 4 (2007): 395-416.
4. Gupta, Naveen, Anurag Singh, and Hocine Cherifi. "Centrality measures for networks with community structure." *Physica A: Statistical Mechanics and its Applications* 452 (2016): 46-59.
5. Fortunato, Santo, and Mark EJ Newman. "20 years of network community detection." *Nature Physics* 18, no. 8 (2022): 848-850.
6. Traag, V. A., L. Waltman, and N. J. Van Eck. "From Louvain to Leiden: guaranteeing well-connected communities." *Scientific Reports* 9 (2019): e5233.
7. Su, Xing, Shan Xue, Fanzhen Liu, Jia Wu, Jian Yang, Chuan Zhou, Wenbin Hu et al. "A comprehensive survey on community detection with deep learning." *IEEE transactions on neural networks and learning systems* 35, no. 4 (2022): 4682-4702.

## 8. Modeling in Complex Networks

Tag: *Social/complex networks*

Complex networks arise in diverse domains, including social systems, biological interactions, communication infrastructure, and information networks. Modeling these systems requires capturing structural heterogeneity, dynamics, and multi-scale interactions. At AIL, we develop mathematical and algorithmic models that explain and predict behavior in complex networks.

Our work integrates graph algorithms, probabilistic modeling, and optimization to study phenomena such as growth processes, diffusion dynamics, robustness, and emergent structure. We aim to balance model expressiveness with analytical tractability, enabling both theoretical insights and scalable computation. Connections to learning on graphs and network inference are also a key focus.

By grounding network modeling in rigorous algorithmic principles, this research provides a foundation for understanding real-world systems and informing applications such as social analysis, recommendation, and decision-making. Modeling in complex networks serves as a unifying theme that connects theory, data, and applications across the lab's research agenda.

### Representative Papers

1. Costa, L. da F., Francisco A. Rodrigues, Gonzalo Travieso, and Paulino Ribeiro Villas Boas. "Characterization of complex networks: A survey of measurements." *Advances in physics* 56, no. 1 (2007): 167-242.
2. Stam, Cornelis J., and Jaap C. Reijneveld. "Graph theoretical analysis of complex networks in the brain." *Nonlinear biomedical physics* 1, no. 1 (2007): 3.
3. Gao, Zhong-Ke, Michael Small, and Juergen Kurths. "Complex network analysis of time series." *Europhysics Letters* 116, no. 5 (2017): 50001.

4. Ji, Peng, Jiachen Ye, Yu Mu, Wei Lin, Yang Tian, Chittaranjan Hens, Matjaž Perc, Yang Tang, Jie Sun, and Jürgen Kurths. "Signal propagation in complex networks." *Physics reports* 1017 (2023): 1-96.
5. Böttcher, Lucas, and Mason A. Porter. "Complex networks with complex weights." *Physical Review E* 109, no. 2 (2024): 024314.
6. Gupta, Naveen, Anurag Singh, and Hocine Cherifi. "Centrality measures for networks with community structure." *Physica A: Statistical Mechanics and its Applications* 452 (2016): 46-59.