**Different versions of bandits:**

1. Stochastic vs Adversial
2. Structured vs Unstructured arms
3. Finite vs Infinite Arms
4. Stationary vs nonstationary rewards
5. Contextual vs without context
6. Linear vs Nonlinear (contextual bandits)
7. Delayed feedback reward vs instant reward

* Duelling bandits
* Convex bandits
* Budgeted bandits
* Graph feedback
* Lipschitz bandits
* Risk Averse
* Neural bandits
* Pure exploration
* Bayesian bandits
* Partial Monitoring

Proposition of master thesis:

* Historical note Applications
* Brief description with the focus on differences of the existing and possible environments: stochastic, adversial etc.
* Unbiased Reward Estimator, measuring the quality, different measures like VAR etc.
* Thompson sampling with linear payoffs
* Epsilon greedy algorithm
* UCB and extensions with the emphasis on context (LinUCB, AdaUCB, SpectralUCB, Cofine UCB …)
* EXP3 algorithm (+ IX)
* LinREL, SupLinREL
* EXP4 algorithm (+ IX)
* Contextual bandits in the view of Empirical Risk Minimisation
* Contextual bandits in the view of multi-class classification

STOCHASTIC BANDITS WITH LEINEAR RELIZABILIYU ASSUMPTION

KERNELIZED STOCHASTIC CONTEXTUAL BANDITS

ARBITRATY SET OF POLICIES

DATASETS

MovieLens dataset

PAPERS

1. **BOOKS: Lattimor Szepesvari + Silvinskis + Regret Analysis MAREC**
2. **Linked Papers MARZEC**
3. **New Papers MARZEC**
4. **Youtube Videos MARZEC**
5. **Master Thesis KWIECIEN**

**Very brief review of concepts in the not contextual setting.**

**Diving deep into contextual setting literature review.**

**A comparison of algorithms on a synthetic dataset.**

**A comparison of algorithms on a real dataset.**

**Topics in contextual bandits. Heavy literature review. Abstracts itp.**

[**https://contextual-bandits.readthedocs.io/en/latest/**](https://contextual-bandits.readthedocs.io/en/latest/)

[**https://yuanz.web.illinois.edu/teaching/IE498fa19/**](https://yuanz.web.illinois.edu/teaching/IE498fa19/)

[**https://proceedings.mlr.press/v15/chu11a.html**](https://proceedings.mlr.press/v15/chu11a.html)

**https://www.ijcai.org/proceedings/2017/0203.pdf**

[**https://arxiv.org/pdf/1605.07139.pdf**](https://arxiv.org/pdf/1605.07139.pdf)

**https://getstream.io/blog/introduction-contextual-bandits/**

**Plan of master thesis:**

1. **Introduction: what is this problem of multi armed bandits and online learning in the highest level possible without any mathematical notation – maybe some short historical context, some very general applications, what is the problem of this thesis the hypothesis or the task (for now it is a comparison of contextual multi-armed bandits), what is done in subsequent chapters and how they are linked.**
2. Brief description with the focus on differences of the existing and possible environments: stochastic, adversial etc. Introduction to notation and fundamental definitions ~ regret, reward, action, environment, history, context etc.
3. **Historical context, applications in greater detail, what multi armed bandits are not, what are contextual multi armed bandits, what is being done in the field, very general very brief.**

**Literature:**

**[1]** Tor Lattimore, Csaba Szepesvári (2020). *Bandit Algorithms*, Cambridge University Press

**[2]** Aleksandrs Slivinks (2019). *Introduction to Multi-Armed Bandits*, Foundations and Trends in Machine Learning, Vol 12, No 1-2, 1-286.

**[3]** Li Zhou (2015). *A survey on Contextual Multi-armed Bandits*, arXive.

**[4]** Chih-Chun Wang, Sanjeev Kulkarani, Vincent Poor (2005). *Bandit problems with side observations*, IEEE Transactions on Automatic Control, 50, 338-355.

**[5]** Sebastien Bubeck, Nicolo Cesa – Bianchi Regret (2012). *Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems*, Now Publihsers

**[6]** Joannés Vermorel, Mehryar Mohri (2005) *Multi-Armed Bandit Algotithms and Empirical Evaluation,* Machine Learning : EMCL 2005, Springer, 437-448.

**[7]**

**[8]**

**[9]**

**[10]**