**ZGŁOSZENIE TEMATU PRACY DYPLOMOWEJ MAGISTERSKIEJ**

rok akademicki 2021/2022

|  |  |
| --- | --- |
| **Imię, nazwisko, tytuł, stopień naukowy** | **Szymon Jaroszewicz prof. dr hab. inż.** |
| e-mail, Instytucja, Wydział (jeśli inne niż PW MiNI) | szymon.jaroszewicz@ipipan.waw.pl |
| **Tytuł zgłaszanej pracy (w jęz.polskim)** | **Optymalizacja w problemie wielorękich kontekstowych bandytów.** |
| **Tytuł zgłaszanej pracy (w jęz.angielskim)** | **Optimization in contextual multi armed bandits.** |
| Kierunek studiów | Matematyka / Mathematics |
| Imię i nazwisko dyplomanta/ki (opcjonalnie) | Hubert Drążkowski |

**Cel pracy:**

Problem wielorękich bandytów (ang. multi-armed bandits) pochodzi z dziedziny uczenia ze wzmocnieniem. To zadanie optymalizacji pewnej funkcji wypłaty przy zadanym ograniczeniu. Jej celem jest wybór strategii zdefiniowanej jako zbiór akcji w procesie sekwencyjnej alokacji z góry ograniczonych zasobów do pewnych, dających nagrodę ramion, tak aby maksymalizować określoną funkcję zysku, będącą skumulowaną wypłatą ze strategii. Celem pracy jest przeanalizowanie algorytmów zajmujących się problemem wielorękich bandytów, w tym wersji kontekstowej.

**Tematyka zgłaszanej pracy:**

W problemie wielorękich bandytów w każdym momencie czasu można dokonać jednej alokacji zasobu do akcji i w zamian otrzymać obserwowalną wypłatę z ramienia. Fundamentalnym problemem zagadnienia jest optymalne rozwiązanie ograniczonych możliwości realizowania konkurujących ze sobą dwóch etapów, eksploracji i eksploatacji. Problem powstaje ze względu na fakt, że niepełna jest informacja o rozkładzie wypłat z ramion. Należy więc z jednej strony gromadzić dane, zaś z drugiej wykorzystywać te dotychczas zebrane do podjęcia jak najlepszej decyzji o alokacji zasobów. Rozszerzenie tematu o kontekst, polega na dodaniu na każdym etapie do historycznych danych informacji o stanie ramion, która wpływa na wypłatę. Pierwsza część pracy powinna zawierać wstęp do klasycznego zagadnienia wielorękich bandytów. Druga część to rozszerzenie problemu do uwzględnienia kontekstu, w tym analiza algorytmów zajmujących się optymalnym rozwiązaniem problemu eksploracji i eksploatacji. Trzecia cześć powinna zawierać nowatorskie rozwiązanie problemu i/lub zaadaptowanie dostępnych metod do rozszerzonego problemu opisanego w drugiej części. W badaniu pojawi się także zastosowanie i symulacja komputerowa na syntetycznych i/lub prawdziwych danych. Zastosowanie opisywanych rozwiązań jest szerokie. Uwzględnia kalibrację portfela inwestycyjnego, badanie leków w testach klinicznych, optymalny dobór strategii marketingowej itp.

**Literatura pomocnicza:**

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

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

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

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

Chih-Chun Wang, Sanjeev Kulkarani, Vincent Poor (2005). *Bandit problems with side observations*, IEEE

Transactions on Automatic Control, 50, 338-355.

Jhon White (2012). *Bandit Algorithms for Website Optimization,* O’Reilly.

Qing Zhao (2019). *Multi-Armed Bandits Theory and Applications to Online Learning in Networks*, Morgan & Claypool Publishers

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

Praca ~~będzie~~ **/ nie będzie** (\*) realizowana przy współudziale lub na zlecenie podmiotów zewnętrznych.

Podmiot zewnętrzny (opcjonalnie):

*Wyrażam zgodę / ~~nie wyrażam zgody~~ (\*) na udostępnienie elektroniczne w sieci Wydziału MiNI pełnego opisu zgłaszanego tematu.*

………………………………………….…………

data i podpis