

Seminarium dyplomowe 1

Maciej Grzenda, PhD

M.Grzenda@mini.pw.edu.pl

<http://www.mini.pw.edu.pl/~grzendam>

Warsaw University of Technology
Faculty of Mathematics and Information Science



**Fundusze
Europejskie**
Wiedza Edukacja Rozwój

**Politechnika
Warszawska**

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn. „Przygotowanie i uruchomienie nowego kierunku studiów na studiach
II stopnia - Inżynieria i Analiza Danych (IAD)”
realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca” współfinansowanego ze
środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego

The objectives of this course

- Support the development of your Master thesis to make them:
 - interesting and valuable
 - high quality documents
- Deliver presentations showing the progress and sharing experience
 - in Polish (during *Diploma Seminar 1*)
 - and in English (during *Diploma Seminar 2*)
- Reduce the time needed to develop Master thesis by:
 - Discussing typical aspects of work preparation
 - Sharing experience arising from Master theses documents developed in the past
 - Refining best practices related to this kind of work
- Mitigate the risk of delays by defining rules of assessment promoting progress in thesis development

Unless otherwise stated, by referring to Seminar, we refer to Seminar 1 in the remainder of this presentation.

Cel finalny: praca magisterska

- Inżynieria i Analiza Danych to podobszar informatyki
- Struktura zgodna ze strukturą pracy informatycznej:
 - Dokument pracy, czyli praca magisterska
 - Płyta z wynikami prac programistycznych i obliczeniowych:
 - kod stworzonego oprogramowania (np. skrypty w R, aplikacje Python, aplikacje Java, pliki konfiguracyjne)
 - Opcjonalnie: maszyna wirtualna, wyniki obliczeń, wykresy, zbiory danych....
- Dokładny zakres zależy od uzgodnień z promotorem, ale dokument pracy i **kod oprogramowania** stanowią część każdej pracy

Typowy układ głównego dokumentu pracy

Praca magisterska

Wstęp	Przegląd	Koncepcja	Wyniki	Podsumowanie
1-3 strony opisu planowanej tematyki	<ol style="list-style-type: none">1. Zazwyczaj ok. 15-30 stron opisu metod, algorytmów, środowisk, bibliotek, wyzwań w dziedzinie pracy, do których będziemy odwoływać się w dalszej części pracy2. Tu wprowadzamy odwołania do większości źródeł literaturowych	<ol style="list-style-type: none">1. Zazwyczaj ok. 10 stron opisu koncepcji rozwiązania problemu podejmowanego w pracy2. Doprecyzowanie celu pracy3. Architektura tworzonego rozwiązania, wykorzystywane komponenty i własne komponenty	<ol style="list-style-type: none">1. Wyniki uzyskane w pracy2. W przypadku użycia metod uczenia maszynowego, analiza poziomu błędów3. Również dyskusja wyników, zalet i wad analizowanych metod	2-4 strony podsumowania uzyskanych wyników

Opcja 1: zalecana minimalna zawartość pracy mgr w KM4

Praca magisterska

Wstęp	Przegląd	Koncepcja	Wyniki	Podsumowanie
1-3 strony opisu planowanej tematyki	<ol style="list-style-type: none">1. Zazwyczaj ok. 15-30 stron opisu metod, algorytmów, środowisk, bibliotek, wyzwań w dziedzinie pracy, do których będziemy odwoływać się w dalszej części pracy2. Tu wprowadzamy odwołania do większości źródeł literaturowych	<ol style="list-style-type: none">1. Zazwyczaj ok. 10 stron opisu koncepcji rozwiązania problemu podejmowanego w pracy2. Doprecyzowanie celu pracy3. Architektura tworzonego rozwiązania, wykorzystywane komponenty i własne komponenty	<ol style="list-style-type: none">1. Wyniki uzyskane w pracy2. W przypadku użycia metod uczenia maszynowego, analiza poziomu błędów3. Również dyskusja wyników, zalet i wad analizowanych metod	2-4 strony podsumowania uzyskanych wyników

SEMINARIUM 1

Opcja 2: drugi wariant minimalnej postaci pracy mgr w KM4

Praca magisterska

Wstęp	Przegląd	Koncepcja	Wyniki	Podsumowanie
<p>1-3 strony opisu planowanej tematyki</p> <p>SEM. I</p>	<p>1. Zazwyczaj ok. 15-30 stron opisu metod, algorytmów, środowisk, bibliotek, wyzwań dziedzinie pracy, do których będziemy odwoływać się w dalszej części pracy</p> <p>2. Tu wprowadzamy odwołania do większości źródeł literaturowych</p>	<p>1. Zazwyczaj ok. 10 stron opisu koncepcji rozwiązania problemu podejmowanego w pracy</p> <p>2. Doprecyzowanie celu pracy</p> <p>3. Architektura tworzonego rozwiązania, wykorzystywane komponenty i własne komponenty</p> <p>SEM. I</p>	<p>1. Wyniki uzyskane w pracy</p> <p>2. W przypadku użycia metod uczenia</p> <p>3. Również dyskusja wyników, zalet i wad analizowanych metod</p> <p>SEM. I</p>	<p>2-4 strony podsumowania uzyskanych wyników</p>

Część przeglądowa

- Przegląd pojęć, metod, algorytmów, ale również np. środowisk, bibliotek i referencyjnych repozytoriów danych do których będziemy odwoływać się w pracy
- Wynik m.in. przeglądu światowej literatury, zarówno monografii, jak i artykułów naukowych i opracowań technicznych
- Tu wprowadzamy odwołania do 85-100% źródeł literaturowych
- W tej części nie wprowadzamy jeszcze opisu własnych rozwiązań
- Identyfikujemy różnice pomiędzy opisywanymi metodami i ich ograniczenia – czyli punkt wyjścia dla stworzenia własnych rozwiązań
- Celem nie jest drobiazgowy opis popularnych metod np. metody propagacji wstecznej. Dobry przegląd przedstawi cechy i założenia algorytmu oraz odwoła się do literatury, w której znajdziemy szczegółowe informacje
- Po opracowaniu tej części, w kolejnych rozdziałach odwołujemy się raczej do części przeglądowej niż zewnętrznych źródeł

**W artykułach naukowych taka część często jest obecna jako rozdział tytułowany
Related works**

Część koncepcyjna

- Celem jest przedstawienie koncepcji i szczegółów technicznych metod, algorytmów, komponentów i systemów tworzonych w pracy
- Pierwsza część pracy, w której koncentrujemy się na własnych rozwiązaniach, czyli pomyśle rozwiązania postawionych zagadnień, w tym np.:
 - Cechach tworzonych metod uczenia maszynowego i ich związkiem z istniejącymi metodami (opisanymi w przeglądzie)
 - Architekturze tworzonego systemu analizy danych, w tym wykorzystaniu istniejących środowisk (opisanych w przeglądzie)
 - Użyciu istniejących bibliotek i zaproponowaniu własnych komponentów
 - Użyciu języków oprogramowania i środowisk

1. **Naturalne jest wykorzystanie gotowych bibliotek, komponentów i systemów. Konieczne jest natomiast precyzyjne udokumentowanie w tekście pracy własnych i istniejących komponentów rozwiązania - by czytelnik mógł jednoznacznie stwierdzić, co stanowi wkład autora.**
2. **Uwaga: możliwe jest tworzenie nowych projektów i rozwijanie istniejących projektów otwartego oprogramowania (ang. open source) w ramach pracy dyplomowej**

Materiały – zagadnienia formalne

- Zasoby strony <https://ww2.mini.pw.edu.pl/studia/dziekanat/dla-dyplomantow-i-promotorow/>, w tym m.in.:
 - Szczegółowe zasady prowadzenia prac dyplomowych i egzaminów dyplomowych na kierunku Inżynieria i Analiza Danych https://ww2.mini.pw.edu.pl/wp-content/uploads/uch_06_2018_79.pdf
 - Informacje ogólne dotyczące obron
 - Wymagania dotyczące formatu pracy dyplomowej
 - Przykładowy szablon pracy dyplomowej w Latex (uwaga: dotyczy pracy inżynierskiej)
- Uzgodnienia z Dziekanatem Wydziału

Materiały – redakcja pracy magisterskiej

1. [OLERUP2013] Agneta Olerup, *Master Thesis in Informatics. Guidelines*, Lund University, 2013,
<https://portal.research.lu.se/ws/files/5885007/3694008.pdf> (*)
2. Ziółkowska, M. ; Okulewicz, M.; Hildebrand, K. (Red)., *Poradnik pisanie pracy dyplomowej*, Samorząd Studentów Politechniki Warszawskiej, 2009,
<http://bcpw.bg.pw.edu.pl/dlibra/docmetadata?id=1524> (*)
3. Zasady przygotowania prac do zgłoszenia w czasopismach np. dla czasopisma *Data Mining and Knowledge Discovery*
https://www.springer.com/computer/database+management+%26;+information+retrieval/journal/10618?print_view=true&detailsPage=pltc_i_2554736

(*) Tytuły dostępne w Bibliotece Głównej PW lub poprzez <http://bg.pw.edu.pl/>
W kwestiach formalnych (np. terminy, opłaty) decydujące są aktualne regulacje Wydziału.

Wymogi związane z obecnością na zajęciach

- Obecność na zajęciach jest obowiązkowa
- Nieobecności bez formalnego uzasadnienia (w postaci zwolnienia lekarskiego lub np. dokumentu potwierdzającego udział w wymianie akademickiej np. Erasmus+):
 - 1-2 nieobecności: bez wpływu na ocenę
 - 3 nieobecności: redukcja oceny
 - 4 nieobecności i więcej: brak zaliczenia przedmiotu
- Dopuszczalna jest zatem dwukrotna nieobecność bez przedstawiania zaświadczenia lekarskiego, ale z wyjątkiem ustalonych terminów referatu oraz prezentacji wyników postępów w pracy
- W wyjątkowych przypadkach, za zgodą prowadzącego i po wcześniejszym uzgodnieniu możliwy jest udział w zajęciach drugiej z grup seminaryjnych

Zasady zaliczeń

- Na ocenę składają się:
 - Ocena referatów wygłaszanych w trakcie seminarium: maks. 50 p.
 - Ocena stopnia zaawansowania pracy magisterskiej: maks. 50 p.
- trzecia nieobecność nieusprawiedliwiona powoduje redukcję oceny o 5 punktów
- Opóźnienie realizacji części przeglądowej o 1 tydzień: redukcja oceny o 7 p.
- Do zaliczenia wymagane są:
 - co najmniej 25 punktów łącznie z referatów (konieczne jest wygłoszenie wszystkich zaplanowanych referatów)
 - co najmniej 25 punktów z tytułu zaawansowania pracy magisterskiej
 - Zrealizowanie wymagań wszystkich kamieni milowych KM1-KM5
- Ocena końcowa zależy od łącznej liczby punktów:

– 51-60	=>	3.0
– 61-70	=>	3.5
– 71-80	=>	4.0
– 81-90	=>	4.5
– 91-100	=>	5.0

Liczba referatów a zasady zaliczeń

- Zgodnie z wcześniejszą informacją, na ocenę składa się m.in. ocena referatów wygłaszanych w trakcie seminarium (maks. 50 p.)
- W zależności od liczności grupy (i tym samym czasu dostępnego na wygłoszenie referatów) realizowany jest jeden z dwóch wariantów:
 - Wariant A: 1 referat z kierunku studiów + 1 referat z tematyki pracy, o wymogach opisanych w dalszej części prezentacji jako KM1 i KM3
 - Wariant B: 1 referat z kierunku studiów + 2 referaty z tematyki pracy, o wymogach opisanych w dalszej części prezentacji jako KM1, KM3 i KM5
- W przypadku realizacji zajęć w wariancie A, suma punktów uzyskanych z obu referatów jest mnożona przez (50/30) w celu zachowania proporcji wpływu oceny referatów i zaawansowania pracy dyplomowej na ocenę finalną

Możliwość poprawy uzyskanych wyników

- Z przyczyn logistycznych nie jest możliwe ponowne przedstawianie referatów
- W przypadku przedstawienia referatów w wyznaczonych terminach i chęci poprawy uzyskanych ocen, możliwa jest poprawa oceny na podstawie przedstawienia w terminie do ostatniego terminu seminarium w danej grupie znacznie poszerzonej wersji pracy magisterskiej
- Nowa wersja pracy magisterskiej musi:
 - znacznie wykraczać poza wymogi KM4 (wymogi dla części przeglądowej pracy opisane w dalszej części pracy)
 - być w pełni dopracowana pod względem edytorskim
 - być zatwierdzona przez promotora pracy

Harmonogram prac

		Tydzień														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
KM1	Wybór i/lub uzgodnienie zgłoszenia tematu pracy															
KM2	Referat: temat kierunkowy															
KM3	Referat: założenia i wstępne wyniki prac															
KM4	Przygotowanie części przeglądowej pracy															
	Przygotowanie roboczej wersji części koncepcyjnej pracy															
KM5	Referat: wyniki części przeglądowej i stan części koncepcyjnej															

Gotowe pierwsze rozdziały pracy

Częściowo równolegle realizowana praca nad koncepcją własnych rozwiązań, w tym architekturą rozwiązania, stosowanymi i tworzonymi metodami, danymi, metodami oceny wyników

Referat prezentujący stan zaawansowania części koncepcyjnej pracy

W przypadku studentów realizujących cały semestr studiów poza Polską, ocena w całości bazuje na przeskalowanych punktach z tytułu realizacji KM4

Seminarium: kamienie milowe

	Produkt	Termin złożenia	Punkty (maks.)
KM1: wybór jednego z gotowych lub uzgodnienie z promotorem dedykowanego zgłoszenia tematu pracy magisterskiej	Deklaracja wyboru tematu w dziekanacie	Koniec VI tygodnia semestru	Nie dotyczy
KM2: przygotowanie i przedstawienie prezentacji z jednego z tematów kierunkowych	Prezentacja na seminarium	Koniec VI tygodnia semestru	10
KM3: prezentacja pierwszego referatu z tematyki pracy dyplomowej	Prezentacja i raport wstępny z projektu	Koniec X tygodnia semestru	20
KM4: przygotowanie części przeglądowej pracy magisterskiej	Dokument tekstu pracy	Koniec X tygodnia semestru	50
KM5: prezentacja drugiego referatu z tematyki pracy dyplomowej (*)	Prezentacja	Koniec ost. tyg. semestru	20

(*) w zależności od liczności grupy, mogą być realizowane 1 lub 2 referaty z pracy.

KM1 – wybór tematu pracy

Celem jest dokonanie wyboru tematu pracy magisterskiej poprzez:

- Wybór jednego z gotowych tematów pracy, zaakceptowanych przez Komisję Programową kierunku IAD
- Lub:
 - Uzgodnienie z jednym z potencjalnych promotorów zgłoszenia tematu pracy magisterskiej i jego zgłoszenie przez promotora do Komisji Programowej kierunku jako dedykowanego dla danej osoby
 - zatwierdzenie tego zgłoszenia przez Komisję Programową kierunku IAD

Ze względu na konieczność formalnej akceptacji zgłoszenia tematu, zgłoszenie takie powinno zostać przedstawione do akceptacji Komisji Programowej do końca czwartego tygodnia semestru

KM1 – promotorzy i współpraca z przedsiębiorstwami

- Promotorem może być osoba ze stopniem naukowym co najmniej doktora, specjalizująca się w zagadnieniach kierunku Inżynieria i Analiza Danych
- Promotor nie musi być pracownikiem Wydziału, ani uczelni, ale musi być nauczycielem akademickim
- Temat pracy może być zgłoszony wspólnie przez promotora głównego – nauczyciela akademickiego i zewnętrznego eksperta np. eksperta z przedsiębiorstwa
- Promotor może być opiekunem ograniczonej liczby prac magisterskich (maks. czterech) na kierunku IAD
- Ze względu na konieczność formalnej akceptacji zgłoszenia tematu, zgłoszenie takie powinno zostać przedstawione do akceptacji Komisji Programowej do końca czwartego tygodnia semestru

KM2: referat kierunkowy

Celem jest przedstawienie na seminarium prezentacji spełniającej warunki:

- Jeden z wylosowanych tematów kierunkowych prezentowanych w wylosowanym terminie
- Istnieje możliwość zmiany wylosowanego tematu na temat związany z wybranym tematem pracy magisterskiej – tylko po osiągnięciu KM1 i co najmniej tydzień przed prezentacją
- Czas referatu:
 - 15-17 minut prezentacji
 - 5 minut na pytania i dyskusję

1. Czas prezentacji jest ściśle ograniczony, ze względu na konieczność zapewnienia czasu dla wszystkich referatów
2. Czas prezentacji w trakcie obrony pracy magisterskiej również jest ograniczony, a jego przekroczenie negatywnie wpływa na ocenę
3. Czas wielu prezentacji jest ograniczony, nawet poniżej 5 minut dla tzw. *pitch deck*
<https://www.forbes.com/sites/alejandrocremades/2018/03/02/how-to-create-a-pitch-deck/#347889ae56c0>

KM3 Referat: założenia i wstępne wyniki prac

Celem jest przedstawienie na seminarium prezentacji spełniające warunki:

- Dyskusja tematu pracy magisterskiej i wyników dotąd wykonanych prac: kluczowe wyzwania, zasoby danych do wykorzystania, kluczowe planowane efekty np. nowe metody analizy danych i określenie ich unikalnych cech, stan zaawansowania przeglądu literatury
- Termin referatu jest przesunięty o 4 tygodnie w stosunku do referatu pierwszego
- Czas referatu:
 - 15-17 minut prezentacji
 - 5 minut na pytania i dyskusję

Ten referat przedstawia wyniki wykonanych prac. Częścią referatu jest zaprezentowanie dokumentu pracy magisterskiej w jego aktualnym kształcie. Dokument ten powinien zostać przesłany do prowadzącego przed seminarium. Zaawansowanie dokumentu odpowiada etapowi semestru.

KM4 Opcja 1: Przygotowanie części przeglądowej pracy

W tej opcji, w terminie przewidzianym dla KM4 należy stworzyć część przeglądową pracy:

- Gotowy jest dokument pracy (w tym nastąpił wybór środowiska np. Latex, Ms Word) z następującymi rozdziałami:
 - Wstęp
 - Rozdział/rozdziały przeglądu zagadnień i literatury kluczowej dla tematu pracy
- W/w rozdziały są kompletne i w pełni dopracowane pod względem edytorskim oraz zaakceptowane przez promotora

1. Podział części przeglądowej na rozdziały i podrozdziały oraz jej treść jest wynikiem uzgodnień z promotorem
2. Z punktu widzenia układu pracy kluczowe jest rozdzielenie prezentacji aktualnego stanu dziedziny pracy (część przeglądowa pracy) i własnych koncepcji, rozwiązań oraz uzyskiwanych wyników (część koncepcyjna i wyniki pracy).

KM4 Opcja 2: Przygotowanie fragmentu części przeglądowej oraz pozostałej części pracy

W tej opcji, w terminie przewidzianym dla KM4 należy stworzyć fragment części przeglądowej pracy oraz co najmniej jednej z kolejnych:

- Gotowy jest dokument pracy (w tym nastąpił wybór środowiska np. Latex, Ms Word) z następującymi rozdziałami:
 - Wstęp
 - Pierwsza część rozdziału/rozdziałów przeglądu zagadnień i literatury kluczowej dla tematu pracy
 - Pierwsza część wybranych pozostałych części pracy np. części koncepcyjnej lub opisu zbiorów danych oraz wyników ich wstępnej analizy
- Przygotowane podrozdziały są kompletne i w pełni dopracowane pod względem edytorskim oraz zaakceptowane przez promotora

1. Podział pracy na rozdziały i podrozdziały oraz ich treść jest wynikiem uzgodnień z promotorem

2. W tej opcji:

- powinny powstać finalne podrozdziały części przeglądowej oraz co najmniej jednej z kolejnych części pracy
- pierwsze podrozdziały części przeglądowej (definicje, kluczowe algorytmy)
- Wstępne prace analityczne i/lub implementacyjne również są opisane w treści pracy. Ten fragment pracy również jest w pełni dopracowany i uzgodniony z promotorem.

KM4: uwagi

- Osiągnięcie KM4 może być zrealizowane w opcji 1 lub opcji 2, opisanej powyżej
- Stworzenie części przeglądowej nie wyklucza dalszych prac nad jej doskonaleniem – nie dążymy do modelu waterfall
- Zalecana jest równoległa praca nad częścią koncepcyjną pracy i implementacją własnych rozwiązań
- Rozpoczęcie implementacji przed częścią przeglądową to jednak bardzo wysokie ryzyko:
 - braku prawidłowego uwzględnienia innych metod, czy też faktu istnienia gotowych bibliotek
 - konieczności przebudowy części koncepcyjnej i implementacji po opóźnionym przygotowaniu części przeglądowej

1. Potwierdzenie osiągnięcia KM4 powinna stanowić pisemna pozytywna opinia promotora
2. Należy przewidzieć czas na ok. dwie iteracje opracowania treści rozdziałów i weryfikacji dokumentu pracy przez promotora
3. Klasyczny efekt KM4: ok. 15 -30 stron tekstu rozdziałów z odwołaniami do literatury
4. Kluczowe wymogi i usterki edytorskie – w kolejnej prezentacji




KM5 Referat: wyniki części przeglądowej i koncepcyjnej

Celem jest przedstawienie na seminarium następującej prezentacji:

- Syntetyczne podsumowanie zrealizowanej części przeglądowej – 3-4 slajdy podsumowania części przeglądowej – liczba i rodzaj wykorzystanych źródeł, kluczowe metody, systemy, biblioteki, wnioski
- Pozostała część: podsumowanie stanu pracy mgr, w tym części koncepcyjnej i szczegółowych założeń dla własnych prac, wynikających z części przeglądowej
- Termin referatu jest przesunięty o ok. 4 tygodnie w stosunku do referatu drugiego
- Czas referatu:
 - 15-17 minut prezentacji
 - 5 minut na pytania i dyskusję

Uwaga: ten referat jest realizowany w zależności od liczności grupy. Zaliczenie może też bazować wyłącznie na KM1 i KM3, czyli jednym referacie kierunkowym i jednym z tematyki pracy.

Efekt finalny: kryteria oceny pracy magisterskiej

Kluczowe kryteria oceny	
Zgodność tytułu pracy dyplomowej z jej treścią.	
Wartość merytoryczna pracy (identyfikacja problemu, sformułowanie celu, dobór i sposób wykorzystania narzędzi, rozwiązanie zadania badawczego/ projektowego/ technologicznego/ organizacyjnego).	 KM3, KM5
Analiza literaturowa, dobór i sposób wykorzystania źródeł.	
Trafność i spójność wniosków (krytyczna analiza osiągniętych wyników w odniesieniu do stanu wiedzy , możliwości dalszych kierunków badań).	 KM4
Układ i redakcja pracy (struktura formalna, przejrzystość, staranność edytorska, poprawność języka, wykorzystanie materiału ilustracyjnego).	 KM4
Inne uwagi.	Pogrubiona czcionka – efekty planowane w trakcie pierwszego semestru Seminarium

Praca dyplomowa w języku angielskim

- Istnieje możliwość przygotowania pracy dyplomowej w języku angielskim za zgodą Prodziekana ds. Nauczania
- Konieczne jest złożenie podania o zgodę, zawierającego uzasadnienie
- Możliwe uzasadnienie: planowane zgłoszenie do publikacji artykułu w języku angielskim np. pracy na konferencję lub do czasopisma, bazującej na wynikach pracy dyplomowej

1. Przygotowanie pracy do publikacji w obszarze Data Science jest szczególnie możliwe i zalecane
2. Typowo prace dotyczą nowych metod analizy danych, w tym metod dedykowanych dla konkretnych zastosowań lub powiązanych zagadnień np. metod wyboru danych testowych, redukcji wymiarowości, przetwarzania tekstu,....
3. Dorobek publikacyjny jest ważnym elementem oceny kandydatów na studia doktoranckie, prezentowanym przez wielu najlepszych kandydatów

Pytania związane z seminarium



**Pytania związane z zawartością merytoryczną pracy
proszę kierować do promotorów**

**W kwestiach związanych z Seminarium, zapraszam do
pytań, w tym na konsultacje**

Aktualne terminy na stronie:

<http://mini.pw.edu.pl/~grzendam/pl/dydaktyka.html>

Zapraszam!

Załącznik 1 Tematy referatów kierunkowych

Relacyjne bazy danych

Grafy

Drzewa

Algorytmy sortujące

Klasyfikacja

Regresja

Uczenie nienadzorowane

Metody oceny klasyfikatorów

Składowanie danych w platformach

Big Data

Platformy NoSQL

Indeksy: bazy relacyjne i platformy NoSQL

Metody głębokiego uczenia

Metoda MapReduce

Uczenie w trybie wsadowym i online

Las losowy

Sieci neuronowe typu MLP

Naiwny klasyfikator Bayesa

Drzewa decyzyjne

Przetwarzanie chmurowe

Wzorzec sharding w platformach Big Data

Redukcja wymiarowości

Środowiska przetwarzania strumieni danych

Uczenie ze wzmocnieniem

Sieci rekurencyjne

Hurtownie danych

Skalowanie wydajności w przetwarzaniu chmurowym

Analiza skupień

Bagging i boosting

Metoda PCA

Metoda SVM

Braki w danych a uczenie maszynowe

Podstawy uczenia maszynowego



**Fundusze
Europejskie**
Wiedza Edukacja Rozwój

**Politechnika
Warszawska**

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn. „Przygotowanie i uruchomienie nowego kierunku studiów na studiach
II stopnia - Inżynieria i Analiza Danych (IAD)”
realizowane jest w ramach projektu
„NERW PW. Nauka – Edukacja – Rozwój – Współpraca” współfinansowanego ze
środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego