

**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI,  
INFORMATYKI I INŻYNIERII BIOMEDYCZNEJ**

KATEDRA AUTOMATYKI

Praca dyplomowa inżynierska

*Oparty na biklasteryzacji system rekomendacji   
z wykorzystaniem hurtowni danych  
Biclustering-based recommendation system   
for a data warehouse*

Autor: *Wojciech Torba*

Kierunek studiów: Automatyka i Robotyka

Opiekun pracy: *dr Patryk Orzechowski*

Kraków, 2016

*Uprzedzony o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631   
z późn. zm.): „ Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystyczne wykonanie albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także uprzedzony o odpowiedzialności dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t.j. Dz. U. z 2012 r. poz. 572, z późn. zm.) „Za naruszenie przepisów obowiązujących   
w uczelni oraz za czyny uchybiające godności studenta student ponosi odpowiedzialność dyscyplinarną przed komisją dyscyplinarną albo przed sądem koleżeńskim samorządu studenckiego, zwanym dalej „sądem koleżeńskim”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i że nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.*

Spis treści

[1. WSTĘP 4](#_Toc455335024)

[1.1 WPROWADZENIE 4](#_Toc455335025)

[1.2 CEL 4](#_Toc455335026)

# WSTĘP

## 1.1 WPROWADZENIE

Obecnie ilość danych jakie należy przetworzyć aby podjąć decyzje z najróżniejszych dziedzin, czy to dotyczących nauki, biznesu, czy codzienności rośnie lawinowo. W procesie analizy danych wspieramy się, często nieświadomie, takimi rozwiązaniami jak hurtownie danych, narzędzia Business Intelligence czy coraz popularniesze Big Data.. Ich głównym zadaniem jest przetworzenie danych w pomocne nam informacje, a te w wiedzę. W procesie tym nierzadko zakres danych ograniczany jest do tych, które uznane zostaną za najbardziej dla nas przydatne, a dzieje się tak, aby uniknąć przeciążenia informacją. Często dzieje się to niejawnie na podstawie naszych wcześniejszych wyborów lub wyborów większej grupy o wspólnych cechach, do której możemy zostać zakwalifikowani.

Metody klasyfikacjji danych klasteryzacja, biklasteryzacja

Wykorzystanie komercyjne

## 1.2 CEL

Przedmiotem mojej pracy inżynierskiej jest zaprojektowanie hurtownii danych dla ogólnodostępnego zbioru danych, a następnie zaproponowanie i zaimplementowanie systemu rekomendacji z wykorzystaniem techniki dwu-grupowania, czyli biklasteryzacji. Hurtownia powinna wspierać operacje odczytu danych z określonych formatów jak CSV, jak również operacje OLAP agregacji, zwijania, rozwijania, filtrowania czy eksploracji danych. System rekomendacji powinien dla przesłanych danych wejściowych sugerować pozycje zbliżone spośród danych zawartych w hurtownii.