

Nazwa modułu: Rocznik: Kod: Punkty ECTS: Wydział: Poziom studiów: Specjalność: Kierunek: Semestr: Profil kształcenia: Język wykładowy: Forma i tryb studiów: Strona www: Osoba odpowiedzialna: Osoby prowadzące:

Opisy efektów kształcenia dla modułu

Kod EKM	Student, który zaliczył moduł wie/umie/potrafi	Powiązania z EKK	Sposób weryfikacji efektów kształcenia (forma zaliczeń)
Wiedza			
M_W001	Korzystanie z istniejących rozwiązań technologii mowy, ocenianie ich i wybieranie najlepszych w danej sytuacji.	IA1A_W18	Aktywność na zajęciach, Wykonanie ćwiczeń, Wykonanie ćwiczeń laboratoryjnych, Kolokwium, Sprawozdanie, Prezentacja
Umiejętności			
M_U001	Samodzielnie implementować algorytmy wykorzystywane w technologii mowy	IA1A_U09, IA1A_U21	Wykonanie projektu, Wykonanie ćwiczeń laboratoryjnych, Projekt, Sprawozdanie
M_U002	Projektowanie nowych zastosowań technologii mowy.	IA1A_U21	Wykonanie ćwiczeń laboratoryjnych, Kolokwium, Sprawozdanie
Kompetencje społeczne			
M_K001	Prezentować działanie i skuteczność technologii mowy	IA1A_K06	Prezentacja

Treść modułu kształcenia (program wykładów i pozostałych zajęć)

Wykład

1. Wprowadzenie i historia technologii mowy.
2. Fizjologia mowy i elementy fonetyki języka polskiego. Dialekty języka polskiego. Właściwości sygnału mowy.
3. Zbiory nagrań mowy i tekstów. Słowniki komputerowe.
4. Redukcja zakłóceń i przetwarzanie wstępne sygnału mowy.
5. Metody parametryzacji i segmentacji sygnału mowy.
6. Kodowanie, kompresja i transmisja mowy. Synteza mowy.
7. Metody rozpoznawania wzorców stosowane w technologiach mowy.
8. Ukryte modele Markowa. Rozpoznawanie słów izolowanych.
9. Wprowadzenie do oprogramowania HTK.
10. Metryka edycyjna. Rozpoznawanie mowy ciągłej.
11. Rozpoznawanie i identyfikacja mówcy. Emocje w głosie.
12. Syntaktyczne modelowanie języka.
13. Semantyczne modelowanie. Ontologie.
14. Interfejs głosowy i systemy dialogowe.
15. Istniejące systemy i przyszłość technologii mowy.

Ćwiczenia audytoryjne

1. Elementy statystyki matematycznej (Bayes, rozkłady, Gaussiany, itd.).
2. Dyskusje o technologiach mowy na podstawie artykułu Younga o rozpoznawaniu mowy oraz Reguła Bayesa i entropia.
3. Analiza spektrogramów mowy.
4. Zapis fonetyczny i słowniki w technologiach mowy, metody oceny jakości systemów.
5. Sprawdzian.
6. Klasyfikator Bayesa i GMM.
7. HMM bez komputera – alg. Brute Force, alg. Forward-Backward.
8. DTW, HMM bez komputera – Algorytm Viterbiego, reg. optymalności Bellmana
9. Sprawdzian.
10. Algorytm Dijkstry i metryka edycyjna.
11. Model n-gramowy i modele semantyczne.
12. Prezentacje studentów na wybrany temat.
13. Prezentacje studentów na wybrany temat.
14. Przedstawienie propozycji komercjalizacji technologii mowy przez studentów.
15. Sprawdzian.

Ćwiczenia laboratoryjne

1. Rozpoznawanie płci na podstawie głosu z wykorzystaniem gotowych pakietów Matlaba.
2. Zaprojektowanie własnego systemu rozpoznawania mowy o ograniczonym słownictwie w oparciu o HTK.
3. Wykonanie korpusu własnej mowy 1.
4. Wykonanie korpusu własnej mowy 2.
5. Zaimplementowanie konkatacyjnego syntezyatora mowy.
6. Rozszerzenie korpusu mowy i usprawnienie syntezyatora.
7. Implementacja prostego kodera i kompresora mowy 1.
8. Implementacja prostego kodera i kompresora mowy 2.
9. Rozpoznawanie słów izolowanych: MFCC+DTW 1.
10. Rozpoznawanie słów izolowanych: MFCC+DTW 2.
11. Implementacja prostego HMM 1.
12. Implementacja prostego HMM 2.
13. Przygotowanie modelu HMM z wykorzystaniem własnego korpusu.

14. Przetestowanie opracowanego systemu rozpoznawania mowy.

15. Poprawki i usprawnienia własnego systemu rozpoznawania mowy opartego o HTK.

Sposób obliczania oceny końcowej

Na podstawie oceny z ćwiczeń laboratoryjnych i oceny ze sprawdzianów obejmujących materiał wykładów i ćwiczeń audytoryjnych oraz aktywności na zajęciach i przygotowania prezentacji.

Wymagania wstępne i dodatkowe

Student powinien umieć programować oraz posiadać wiedzę o przetwarzaniu sygnałów, statystyce i probablistyce.

Zalecana literatura i pomoce naukowe

1. B. Ziółko, M. Ziółko, "Przetwarzanie mowy". Wydawnictwa AGH, 2011.
2. S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. Woodland, "HTK Book". UK: Cambridge University Engineering Department, 2005.
3. S. Young, "Large vocabulary continuous speech recognition: a review". IEEE Signal Processing Magazine, vol. 13(5), pp. 45-57, 1996.
4. L. Rabiner and B. H. Juang, "Fundamentals of speech recognition". New Jersey: PTR Prentice-Hall, Inc., 1993.
5. D. Jurafsky and J. H. Martin, "Speech and Language Processing, 2nd Edition". New Jersey: Prentice-Hall, Inc., 2008.
6. C. D. Manning, "Foundations of Statistical Natural Language Processing". MIT Press. Cambridge, MA, 1999.
7. J. N. Holmes, "Speech Synthesis and Recognition". London: Taylor and Francis, 2001.

Uwagi

Brak

Nakład pracy studenta (bilans punktów ECTS)

Forma aktywności studenta	Obciążenie studenta
Udział w wykładach	30 godz
Udział w laboratoriach	30 godz
Udział w ćwiczeniach audytoryjnych	24 godz
Samodzielne studiowanie tematyki wykładów	30 godz
Kolokwium zaliczeniowe	6 godz
Przygotowanie wystąpienia/prezentacji	10 godz
Przygotowanie się do zajęć	20 godz
Samodzielne opracowanie sprawozdania	10 godz
Przygotowanie do kolokwium zaliczeniowego	50 godz
Sumaryczne obciążenie pracą studenta	210 godz
Punkty ECTS za moduł	7 ECTS