



Module name: Hardware Acceleration of Artificial Intelligence Algorithms

Academic year: 2019/2020 Code: ZSDA-3-0102-s ECTS credits: 3

Faculty of: Szkoła Doktorska AGH

Field of study: Szkoła Doktorska AGH Specjalty: —

Study level: Third-cycle studies Form and type of study: Full-time studies

Lecture language: Polski i Angielski Profile of education: Academic (A) Semester: 0

Course homepage: —

Responsible teacher: dr hab. inż. Jamro Ernest (jamro@agh.edu.pl)

Module summary

Współczesne algorytmu uczenia maszynowego, w szczególności głębokie sieci neuronowe, wymagają znacznych mocy obliczeniowych, dlatego kluczowe znaczenie w praktycznej implementacji ma użyta platforma sprzętowa. Rozwiązania sprzętowe przedstawione zostaną w odniesieniu do popularnych algorytmów sztucznej inteligencji i architektury najczęściej używanych sieci neuronowych.

Description of learning outcomes for module

MLO code	Student after module completion has the knowledge/ knows how to/is able to	Connections with FLO	Method of learning outcomes verification (form of completion)
Social competence: is able to			
M_K001	Zrozumienie ograniczeń i praktycznych możliwości wykorzystania sztucznej inteligencji w wybranych dziedzinach życia.	SDA3A_K01	Scientific paper
Skills: he can			
M_U001	Umiejętność wyboru platformy sprzętowej dla zadanych parametrów projektowych do realizacji algorytmu sztucznej inteligencji	SDA3A_U01, SDA3A_U03	Project
M_U002	Umiejętność użycia wybranej platformy sprzętowej do algorytmu głębokiego uczenia	SDA3A_U01, SDA3A_U03	Project
Knowledge: he knows and understands			
M_W001	Znajomość współczesnych sprzętowych platform obliczeniowych i ich cech w kontekście zastosowania do implementacji algorytmów sztucznej inteligencji	SDA3A_W02, SDA3A_W01	Scientific paper

Number of hours for each form of classes

Suma	Form of classes										
	Lectures	Auditorium classes	Laboratory classes	Project classes	Conversation seminar	Seminar classes	Practical classes	Fieldwork classes	Workshops	Prace kontrolne i przejściowe	Lektorat
30	14	0	0	8	8	0	0	0	0	0	0

FLO matrix in relation to forms of classes

MLO code	Student after module completion has the knowledge/ knows how to/is able to	Form of classes										
		Lectures	Auditorium classes	Laboratory classes	Project classes	Conversation seminar	Seminar classes	Practical classes	Fieldwork classes	Workshops	Prace kontrolne i przejściowe	Lektorat
Social competence: is able to												
M_K001	Zrozumienie ograniczeń i praktycznych możliwości wykorzystania sztucznej inteligencji w wybranych dziedzinach życia.	-	-	-	-	+	-	-	-	-	-	-
Skills: he can												
M_U001	Umiejętność wyboru platformy sprzętowej dla zadanych parametrów projektowych do realizacji algorytmu sztucznej inteligencji	-	-	-	+	+	-	-	-	-	-	-
M_U002	Umiejętność użycia wybranej platformy sprzętowej do algorytmu głębokiego uczenia	-	-	-	+	-	-	-	-	-	-	-
Knowledge: he knows and understands												
M_W001	Znajomość współczesnych sprzętowych platform obliczeniowych i ich cech w kontekście zastosowania do implementacji algorytmów sztucznej inteligencji	+	-	-	-	-	-	-	-	-	-	-

Student workload (ECTS credits balance)

Student activity form	Student workload
Udział w zajęciach dydaktycznych/praktyka	30 h
Preparation for classes	15 h
przygotowanie projektu, prezentacji, pracy pisemnej, sprawozdania	15 h
Realization of independently performed tasks	15 h
Summary student workload	75 h
Module ECTS credits	3 ECTS

Additional information

Module content

Lectures

Architektura i organizacja komputera, struktura i działanie jednostki centralnej, jednostka stało- i zmiennoprzecinkowa, jednostka wektorowa, hierarchia pamięci, systemy komputerowe obliczeń wielkiej skali, architektury obliczeń masywnie równoległych i ko-procesorów graficznych GPU.

Architektura układów programowalnych FPGA, obliczenia AI w układach FPGA.

Dedykowane procesory sprzętowe do obliczeń AI.

Architektury sieci neuronowych i innych algorytmów AI. Metody ich implementacji w różnych typach procesorów. Optymalizacja architektury sieci neuronowej pod kątem zasobów sprzętowych, np. wpływ szerokości bitowej obliczeń na wydajność oraz na rezultat obliczeń AI. Gotowe biblioteki do obliczeń AI i ich implementacja sprzętowa.

Project classes

Ćwiczenia projektowe dotyczyć będą programowania równoległego na potrzeby algorytmów uczenia maszynowego: obliczenia wielkiej skali w środowisku superkomputerowym; obliczenia współbieżne SMP przy pomocy OpenMP; obliczenia klastrowe w środowisku MPI; programowanie GPGPU w CUDA i OpenCL; projektowanie struktury FPGA w językach wysokiego poziomu; implementacja sieci neuronowych na architekturach obliczeń równoległych; przykładowa implementacja sieci neuronowej w układach dedykowanych; zastosowanie dedykowanych procesorów sprzętowych w obliczeniach AI.

Conversation seminar

Przykładowe tematy proponowane do dyskusji:

Przykładowe zastosowania algorytmów AI.

Przykłady wykorzystanie architektur dedykowanych w obliczeniach produkcyjnych.

Strategie wyboru architektury sprzętowej; wydajność obliczeń, a zużyte zasoby.

Kierunki dalszej optymalizacji algorytmów AI i architektury sprzętowej.

Teaching methods and techniques:

Lectures: prezentacje w powerponcie, ankieta po wykładzie kahoot

Project classes: zajęcia projektowe, użycie stosownych narzędzi i systemów do projektowania.

Conversation seminar: Nie określono

Warunki i sposób zaliczenia poszczególnych form zajęć, w tym zasady zaliczeń poprawkowych, a także warunki dopuszczenia do egzaminu:

Brak egzaminu. Warunkiem uzyskania zaliczenia jest oddanie wybranego projektu oraz przygotowanie prezentacji i aktywny udział w konwersatorium.

Zasady udziału w poszczególnych zajęciach, ze wskazaniem, czy obecność studenta na zajęciach jest obowiązkowa:

Lectures:

- Attendance is mandatory: No
- Participation rules in classes: dobrowolne

Project classes:

- Attendance is mandatory: Yes
- Participation rules in classes: obowiązkowe

Conversation seminar:

- Attendance is mandatory: Yes
- Participation rules in classes: obowiązkowe

Method of calculating the final grade

Średnia ocen uzyskanych z projektu oraz z konwersatorium.

Sposób i tryb wyrównywania zaległości powstałych wskutek nieobecności studenta na zajęciach:

Projekt można oddać poprzez indywidualne umówienie się z prowadzącym. Konwersatorium można oddać poprzez indywidualną prezentację z prowadzącym.

Prerequisites and additional requirements

Podstawy projektowania w języku C.

Recommended literature and teaching resources

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
2. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.
3. François Chollet, Deep Learning with Python, Manning Publications Co. 2018,
4. Kirk, D. B., & Wen-Mei, W. H. (2016). Programming massively parallel processors: a hands-on approach. Morgan kaufmann.
5. materiały online firmy intel: www.intel.com
6. materiały online firmy nvidia www.nvidia.com
7. materiały online firmy xilinx: www.xilinx.com

Scientific publications of module course instructors related to the topic of the module

- 1) Jamro E., Dąbrowska-Boruch, A., Russek P., Wielgosz M., Wiatr K., Novel architecture for floating point accumulator with cancelation error detection, Bulletin of the Polish Academy of Sciences. Technical Sciences, 2018 vol. 66 no. 5, s. 579-587.
- 2) Russek, P., Jamro, E., Dąbrowska-Boruch, A., & Wiatr, K. (2019). A study of the loops control for reconfigurable computing with OpenCL in the LABS local search problem. The International Journal of High Performance Computing Applications, 1094342019868515.

Additional information

None