

analiza danych

Joanna Szyda - wykład

Marek Sztuka - ćwiczenia

1. Theta

2. Charakterystyka przedmiotu

3. Kontakt

4. Literatura

The screenshot shows the THETA website homepage. At the top, there are navigation links: Home Page, Contact Us, Teaching, Job, and Login. Below this is a header with the text "THETA Statistical Genetics Group Institute of Animal Genetics". A featured article is highlighted with a "JOURNAL ARTICLE" and "EDITOR'S CHOICE" badge. The article title is "Exploring the impact of sequence context on errors in SNP genotype calling with whole genome sequencing data using AI-based autoencoder approach". Below the title, the authors are listed: Krzysztof Kotlarz, Magda Mielczarek, Przemyslaw Biecek, Bernd Guldbrandtsen, and Joanna Szyda. The article is from NAR Genomics and Bioinformatics, Volume 6, Issue 3, September 2024. Below the article, there is a section titled "THETA students at the Symposium of the Polish Bioinformatic Society (PTBI) in Warsaw" and a list of posters presented during the conference.

Home Page Contact Us Teaching Job Login

THETA
Statistical Genetics Group
Institute of Animal Genetics

New one in THETA portfolio. AE published in NAR Genomics and Bioinformatics (Editor's Choice!!!)

JOURNAL ARTICLE EDITOR'S CHOICE

Exploring the impact of sequence context on errors in SNP genotype calling with whole genome sequencing data using AI-based autoencoder approach

Krzysztof Kotlarz, Magda Mielczarek, Przemyslaw Biecek, Bernd Guldbrandtsen, Joanna Szyda

NAR Genomics and Bioinformatics, Volume 6, Issue 3, September 2024, lqae131, <https://doi.org/10.1093/nargab/lqae131>

Published: 24 September 2024 Article history ▾

THETA students at the Symposium of the Polish Bioinformatic Society (PTBI) in Warsaw

Four students: Marek Sztuka, Klaudiusz Tomczyk, Patryk Mierzejewski, Jakub Liu were attending the annual PTBI Symposium, which was held in Warsaw during the days 11-13.09.2024.

Posters presented during the conference included:

1. Application of Neural Network for Common Carp microbiome. Sztuka M., Jakimowicz, M., Kotlarz, K., Mielczarek, M., Naport-Rutkowski, L., Hajduk, P., Szyda J.
2. Analysis of linkage disequilibrium for next generation sequencing data.

Because we like it! 🍌

Staff:

- Joanna Szyda
- Tomasz Suchocki
- Magdalena Frąszczak
- Magda Mielczarek
- Bartosz Czech
- Michalina Jakimowicz
- Barbara Kosińska-Selbi
- Krzysztof Kotlarz
- Jakub Liu
- Dawid Słomian
- Marek Sztuka

SKNBIO
INFORM
A=TYKÓW
UPWROC

Seminars

Feel free to join theta online seminars.
The meeting time is **Wednesday** at **13:30 - 14:30**. Stay tuned and follow the schedule below or request to be added to mailing list.
Contact person: Marek Sztuka.

Winter semester 2024/2025
Seminar schedule:

- 16.10.24 - Magda Mielczarek - Epigenomic Approaches for the
- Diagnosis of Rare Diseases by Beatriz Martinez-Delgado and Maria J. Barrero

projekt THETA → zastosowanie DL do klasyfikacji stawów

materiał



CONTROL

5 ponds

water supplementation
feed supplementation

SETUP 1

5 ponds

water supplementation ▶ A
feed supplementation

SETUP 2

5 ponds

water supplementation ▶ A
feed supplementation

SETUP 3

5 ponds

water supplementation ▶ B
feed supplementation

SETUP 4

5 ponds

water supplementation ▶ B
feed supplementation

projekt THETA → zastosowanie DL do klasyfikacji stawów

materiał



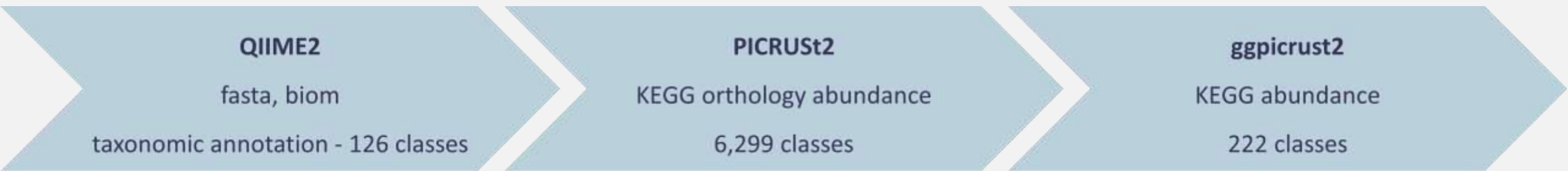
- 16S rRNA



- Woda → 5 próbek / staw x3 czasy
- Sedymnt → 5 próbek / staw
- Jelito ryb → 5 osobników / staw

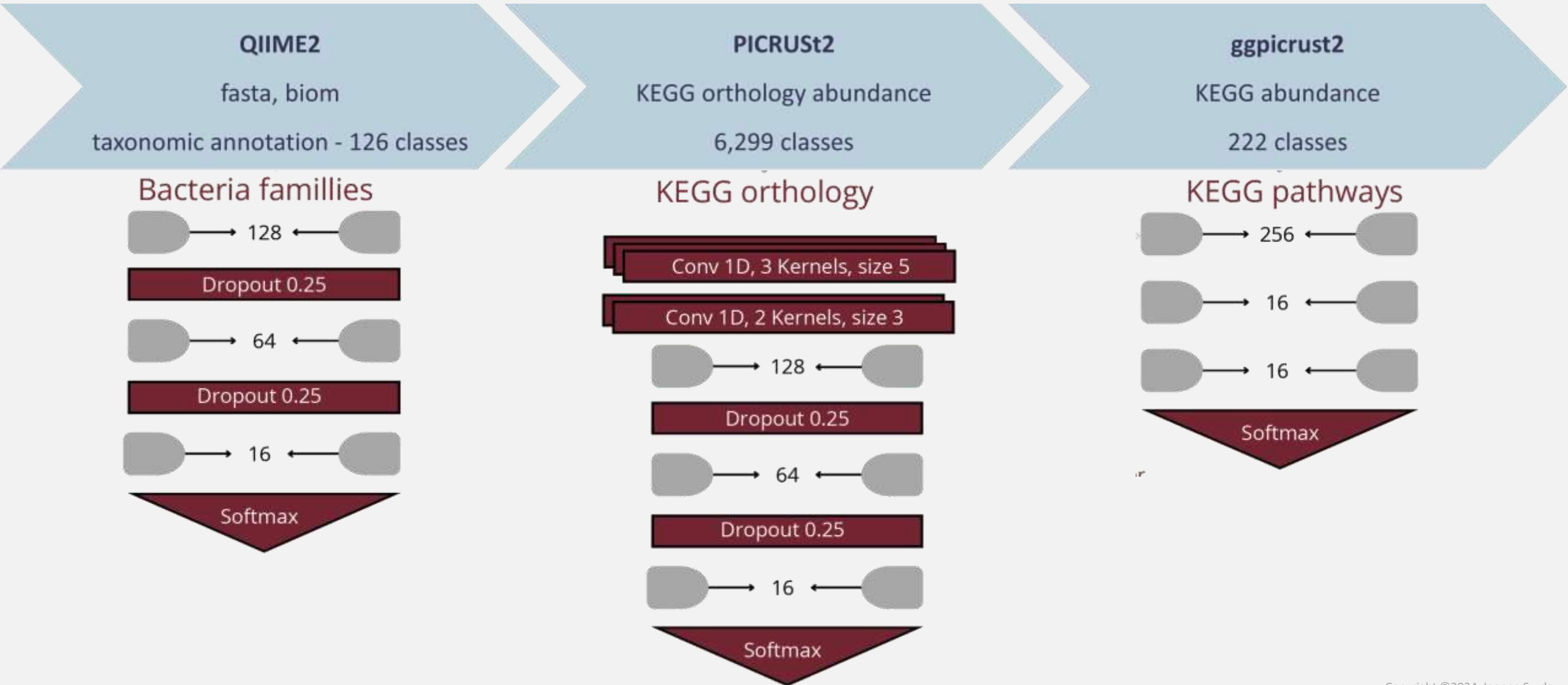
projekt THETA → zastosowanie DL do klasyfikacji stawów

Analiza bioinformatyczna – adnotacja mikrobiologiczna i funkcjonalna



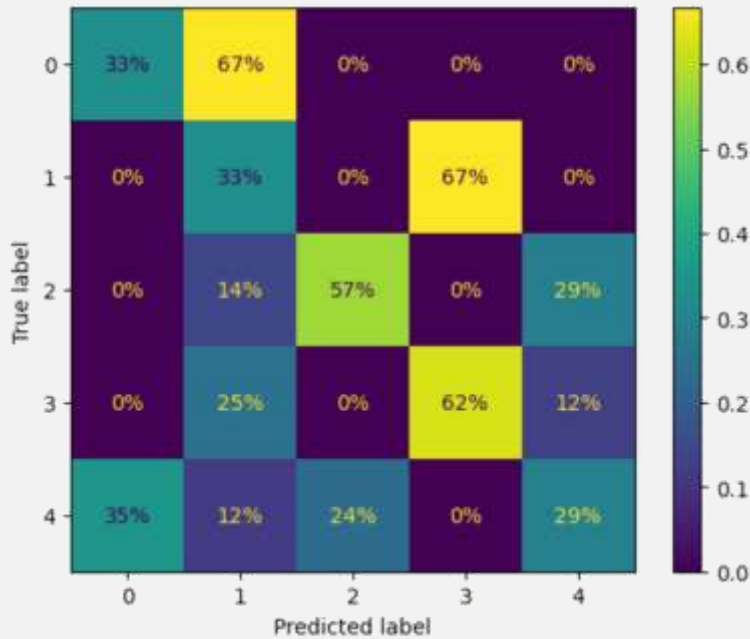
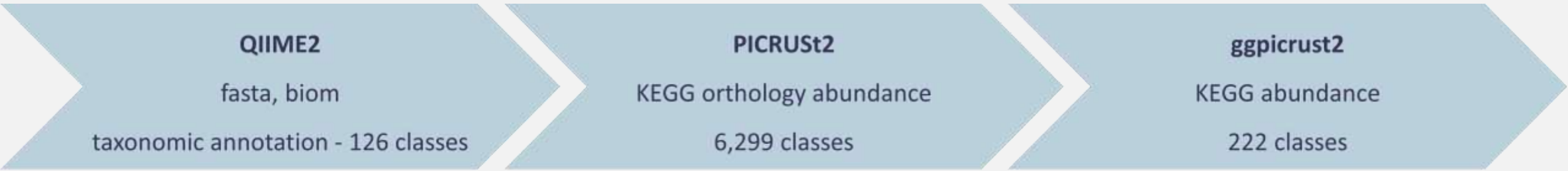
projekt THETA → zastosowanie DL do klasyfikacji stawów

Analiza bioinformatyczna – klasyfikacja stawów DL

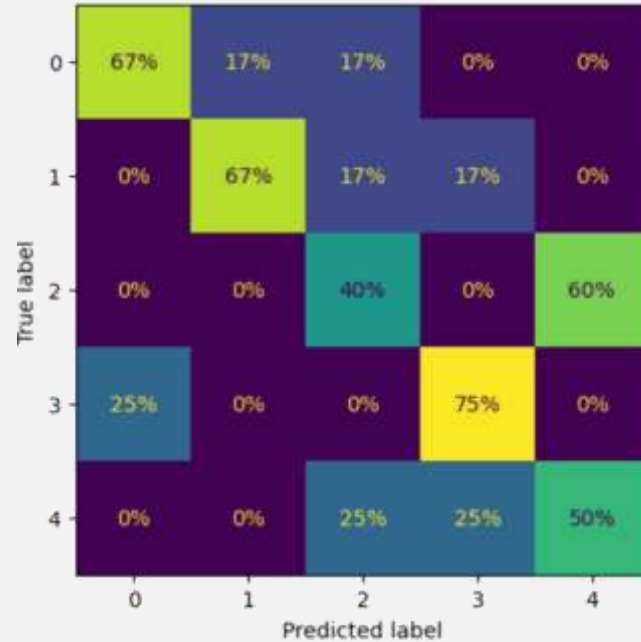


projekt THETA → zastosowanie DL do klasyfikacji stawów

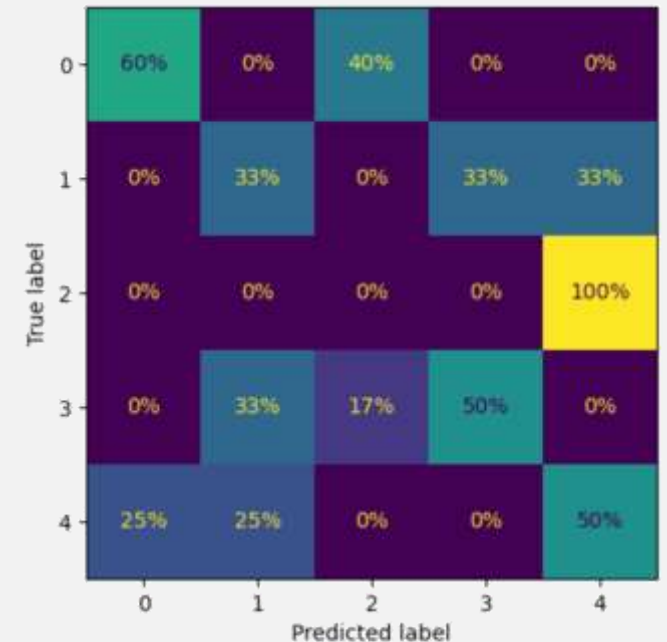
Wyniki – klasyfikacja stawów DL → 5 klas



taxonomy (family)
DNN → ACC=0.43



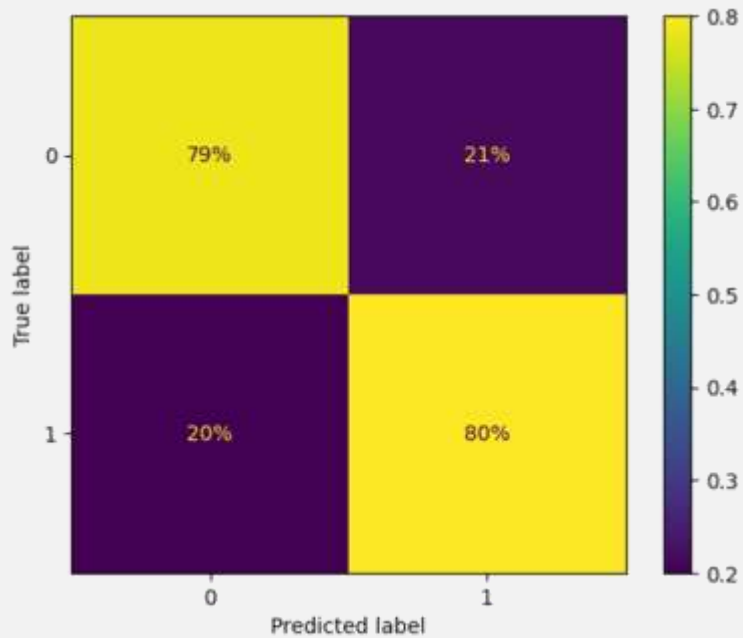
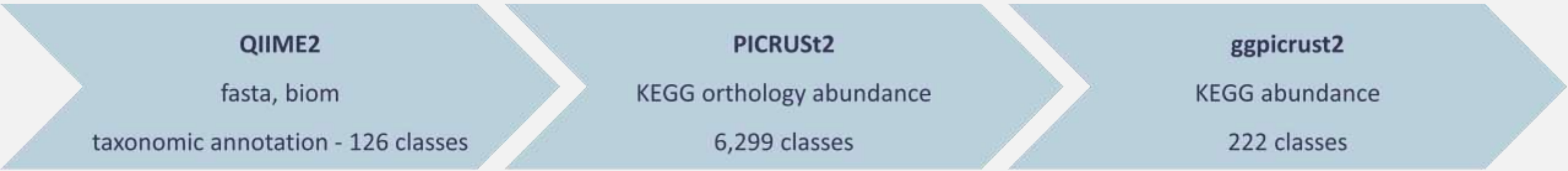
orthology
CNN → ACC=0.64



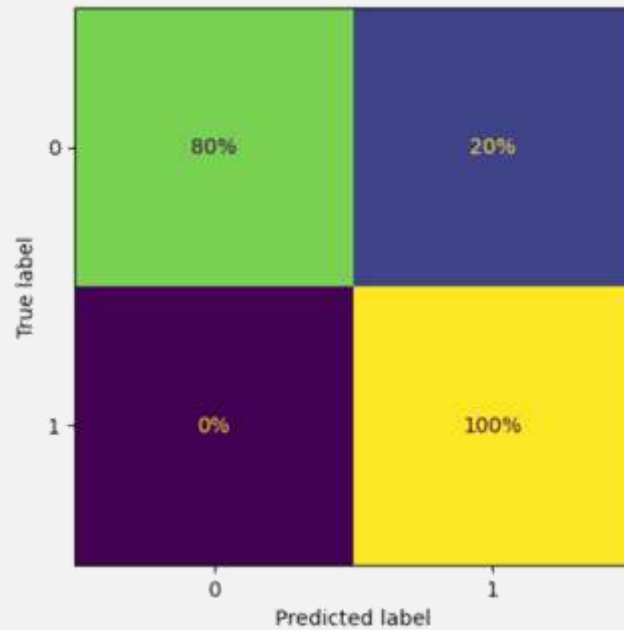
KEGG
DNN → ACC=0.53

projekt THETA → zastosowanie DL do klasyfikacji stawów

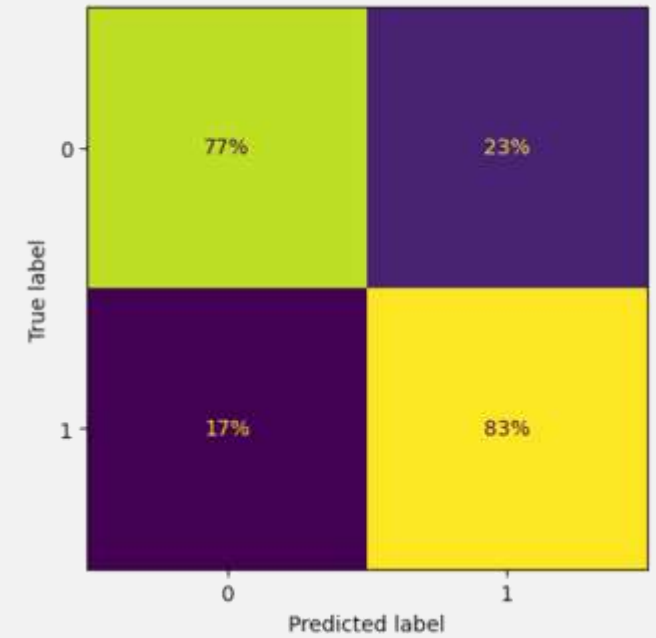
Wyniki – klasyfikacja stawów DL → 2 klasy



taxonomy (family)
DNN → ACC=0.84



orthology
CNN → ACC=0.89



KEGG
DNN → ACC=0.84

1. Theta

2. Charakterystyka przedmiotu

3. Kontakt

4. Literatura

Charakterystyka przedmiotu

- Analiza danych
 - Deep Learning
 - Uczenie głębokie
 - **Keras (Python, R)**
 - TensorFlow (C++, CUDA)
 - → → → PyTorch (Python, C++, CUDA)
- Wykład – teoria, elementy kodu
- Ćwiczenia
 - Praktyczna, samodzielna analiza danych (supervised)
 - Projekt
 - Laptop

przykładowe dane

Constructed from biom file

#OTU ID S8.4.Gut.S8 S8.5.Gut.S8 S8.2.Gut.S8 S8.3.Gut.S8

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Aeromonadaceae 13353.0 12078.0 17717.0 18658.0

d__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae 1438.0 3342.0 1707.0 3280.0

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;__ 193.0 9.0 90.0 19.0

d__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae 183.0 1.0 36.0 2.0

d__Bacteria;p__Firmicutes;c__Bacilli;o__Erysipelotrichales;f__Erysipelotrichaceae 112.0 219.0 173.0 414.0

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Rhodocyclaceae 148.0 26.0 35.0 17.0

d__Bacteria;p__Fusobacteriota;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae 94.0 33.0 69.0 24.0

d__Bacteria;p__Cyanobacteria;c__Cyanobacteriia;o__Cyanobacteriales;f__Phormidiaceae 19.0 162.0 20.0 114.0

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Burkholderiales;f__Chitinibacteraceae 14.0 4.0 25.0 19.0

d__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Carnobacteriaceae 10.0 109.0 69.0 20.0

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Cardiobacteriales;f__Cardiobacteriales 3.0 11.0 14.0 17.0

d__Bacteria;p__Firmicutes;c__Bacilli;o__RsaHf231;f__RsaHf231 0.0 5.0 0.0 20.0

d__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;__;__ 0.0 0.0 87.0 0.0

1. Wykład wstępny
2. Elementy algebry liniowej - przypomnienie
3. Wybrane aspekty obliczeń komputerowych
4. Architektura pojedynczej sieci neuronowej cz.1
5. Architektura pojedynczej sieci neuronowej cz. 2
6. Architektura modelu uczenia głębokiego - wprowadzenie
7. Architektura modelu uczenia głębokiego - rozszerzenia

8. Walidacja modelu

9. Przykłady innych struktur sieci

10. Przykłady własnych badań cz. 1

11. Przykłady własnych badań cz. 2

12. -15. Prezentacje zaliczeniowe

1. Theta

2. Charakterystyka przedmiotu

3. Kontakt

4. Literatura

kontakt

theta.edu.pl/teaching/ → Analiza danych

Home Page Contact Us Teaching Login

THETA
Statistical Genetics Group
Institute of Animal Genetics

Teaching

- Analiza danych
- Bioinformatyka
- Informatyka w selekcji
- Metody statystyczne w biologii
- Modele liniowe i mieszane na przykładzie analizy danych biologicznych
- Podstawy statystycznej analizy danych
- Systemy Informatyczne Wspomagające Hodowlę
- SKN Bioinformatyków

- News
- Alumni
- Collaboration
- Conferences
- Grants
- Projects
- Publications
- Seminars

kontakt

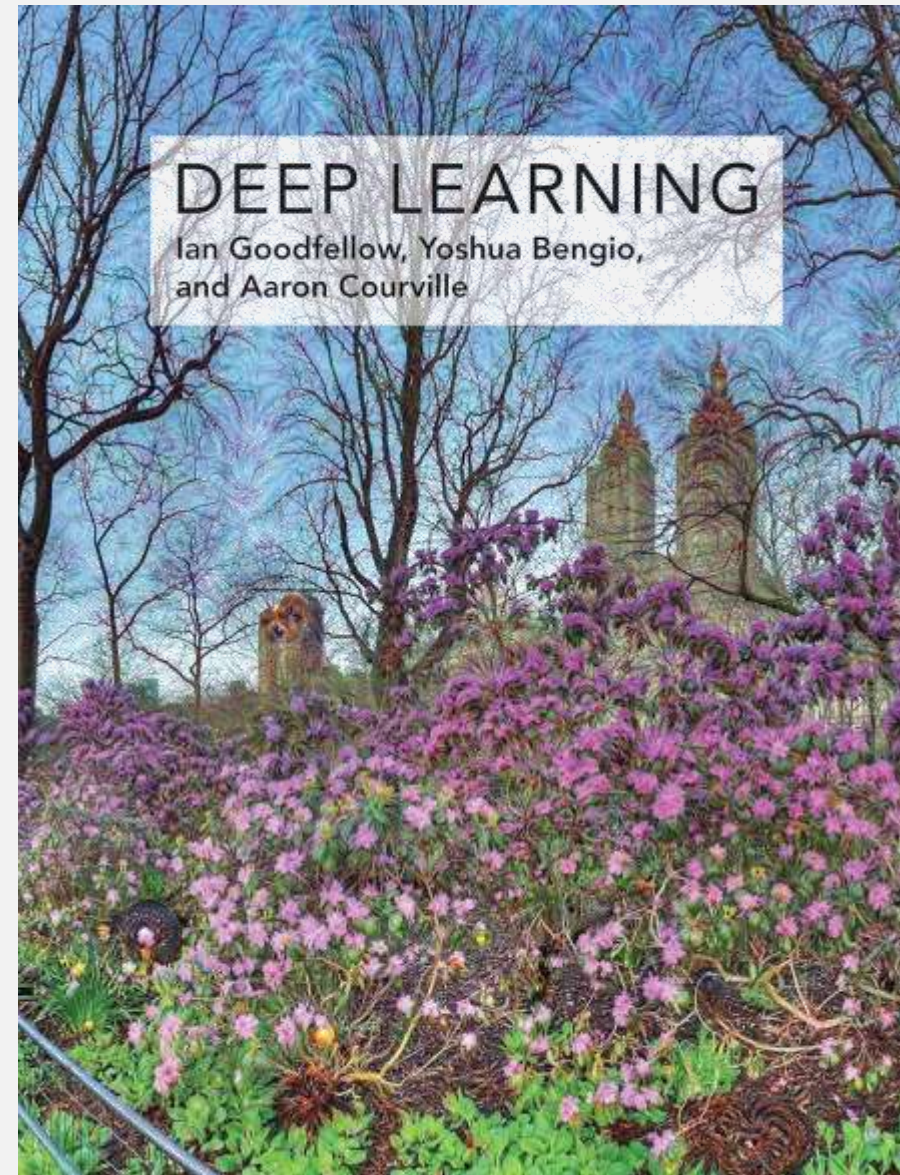
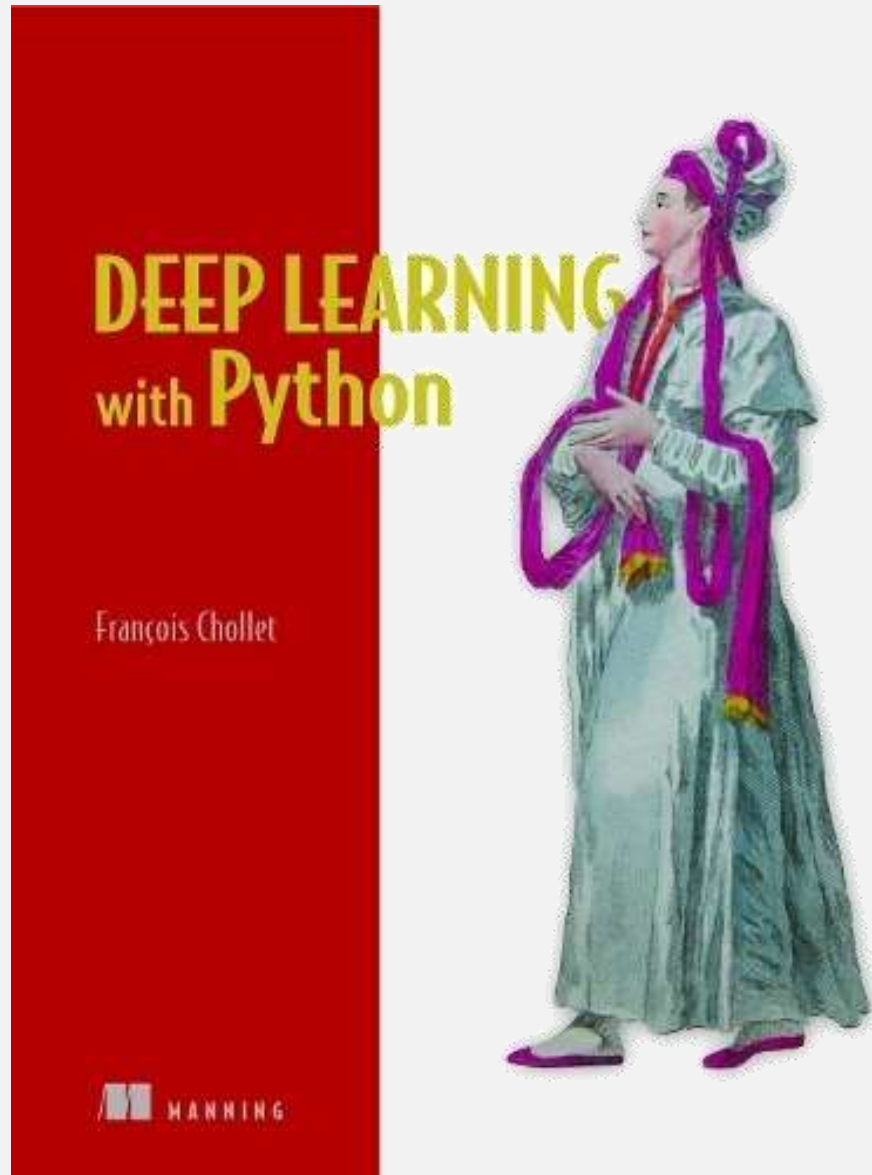
Katedra Genetyki, Kożuchowska 7

Konsultacje

- termin ustalony indywidualnie
- online
- stacjonarne
- Email



-
1. Theta
 2. Charakterystyka przedmiotu
 3. Kontakt
 - 4. Literatura**



1. Theta
2. Charakterystyka przedmiotu
3. Kontakt
4. Literatura

1. Co to jest keras?
2. Jakie dane będziemy analizować ?
3. Jaką metodykę zastosujemy do analizy?
4. Na jakiej podstawie nastąpi zaliczenie przedmiotu?