

Witaj na kursie SAD 2025!

SAD 2025 | Karol Flisikowski | Lato 2025 | PG

O kursie

Ten kurs nauczy Cię, jak myśleć o **danych i modelowaniu dla komputerowych nauk społecznych**.

- Jak jest **formatowany** zbiór danych i czy jest to odpowiedni format dla tego, co chcę zrobić?
- Czy ten zbiór danych jest **reprezentatywny** lub czy odzwierciedla **obiektywną próbę**?
- Jakie **rozważania etyczne** powinienem wziąć pod uwagę podczas uzyskiwania i analizowania danych?
- Jaki **model** jest najbardziej odpowiedni dla tych danych?
- W jaki sposób **projektować i wdrażać** te modele - od [regresji liniowej] (https://en.wikipedia.org/wiki/Linear_regression) do [maszyn wektorów nośnych] (https://en.wikipedia.org/wiki/Support_vector_machine)?

Ten kurs jest jednym z **dwóch podstawowych kursów** z zakresu statystycznej analizy danych, wraz z “Analizą i wizualizacją danych w Pythonie”. Należy pamiętać, że SAD jest warunkiem wstępnym dla tego kursu.

Ocena końcowa

Ocena końcowa wystawiona będzie na podstawie Twojego projektu końcowego.

Nie ma egzaminu końcowego; zobacz syllabus by zobaczyć wszystkie szczegóły nt. kursu.

Obecna edycja

Obecna edycja SAD na studiach podyplomowych to **Lato 2025**. Prowadzącym jest Karol Flisikowski.

Tematy

Omawiane tematy będą obejmować:

- Kwestie etyczne analizy danych: stronniczość, uczciwość, prywatność i inne.
- Porządkowanie i czyszczenie danych (część 1/2).
- Wizualizacja danych (część 1/2).
- Statystyka opisowa.
- Statystyka matematyczna.
- Klasyfikacja (podstawy).
- Podejścia nienadzorowane (podstawy).

Więcej informacji znajdziesz w syllabusie.