

Számítógépes nyelvészet alapozó

Lagos Mátyás
lagosmatyas@gmail.com

2023. ősz

Áttekintés

Ez a kurzus két részből áll: az első rész egy bevezetés a programozásba a Python programozási nyelven, a második részben pedig egy bizonyos típusú számítógépes nyelvészeti modellnek, az n-gram nyelvmodelleknek egy egyszerű változatával fogunk foglalkozni elméleti és gyakorlati szinten. Előzetes programozási ismeretek nem szükségesek, de laptopot mindenkinek kell hozni minden órára.

Követelmények

A félév végi jegyhez minden héten (kivéve a zh-s heteket) házi feladatok megoldásával és két zh megírásával lehet pontokat szerezni. Házi feladatonként 3 pontot és zh-nként 5 pontot lehet szerezni, tehát az összpontszám 40 pont. Az év végi jegy 32 ponttól 5-ös, 28–31,5 pont között 4-es, 24–27,5 pont között 3-as, és 20–23,5 pont között 2-es, de ahhoz hogy valaki 4-est vagy 5-öst tudjon szerezni, a két zh-n összesen legalább 4 pontot el kell érnie.

Tanterv

Óra	Dátum	Téma
1.	09.12.	Előkészülés, a parancssor használata
2.	09.19.	Egy számítógépes program összetétele, stringek, elágazások
3.	09.26.	Ciklusok, listák
4.	10.03.	Szótárak, halmazok
5.	10.10.	A programozás folyamata
6.	10.17.	Összefoglalás, első zh
7.	10.24.	Az n-gram nyelvmodellek elmélete
–	10.31.	Őszi szünet
8.	11.07.	Bigram nyelvmodell programozása 1.
9.	11.14.	Összefoglalás, második zh
10.	11.21.	Bigram nyelvmodell programozása 2.
11.	11.28.	Általánosítás bigram nyelvmodellel 1.
12.	12.05.	Általánosítás bigram nyelvmodellel 2.
13.	12.12.	Szövegfeldolgozás

Ajánlott irodalom

Jurafsky és Martin (2021): *Speech and language processing ch. 3: N-gram language models*
Matthes (2019): *Python crash course*
Zingaro (2021): *Learn to code by solving problems*