

راهنمای واحد درسی مطالعه مستقل

مدرس / مدرسین: دکتر عذرا دائی - دکتر شفیع حبیبی

پیش نیاز یا واحد همزمان: روش تحقیق

تعداد واحد: ۲ نوع واحد: ۲ واحد نظری مقطع: کارشناسی ارشد

تعداد جلسات: ۱۷

زمان برگزاری جلسات در هفته: طبق برنامه آموزش

مکان برگزاری جلسات حضوری: دانشکده مدیریت و اطلاع رسانی پزشکی

هدف کلی و معرفی واحد درسی:

ایجاد کارایی و توانمند سازی دانشجویان در متن کاوی به وسیله ابزار برنامه نویسی پایتون

انتظار می رود فراگیران بعد از گذراندن این دوره بتوانند :

توصیه می شود حداقل ۸۰ درصد اهداف آموزشی براساس آخرین نسخه کوریکولوم مربوطه نوشته شوند و دربرگیرنده اهداف/دانشی، نگرشی و سایکوموتور نیز باشند.

1. معرفی

- a. مقدمه
- b. آشنایی با پایتون
- c. آشنایی با برنامه نویسی و مقدمات آن
- d. آشنایی با کاربردها پایتون مانند زبانهای برنامه نویسی بزرگ LLM, و یادگیری ماشین Machine leaning
- e. NLP چیست؟
- f. کاربردهای NLP
- g. نصب پایتون، NLTK و ...

2. آشنایی با پایتون

- a. معرفی انواع داده ها مانند int, float, str, boolean
- b. مفهوم Assignment و تعریف متغیر انواع متغیر، نام گذاری متغیر
- c. معرفی Operator عملگرها، عملگرهای حسابی و مقایسه‌ای، عملگرهای تشخیص و عضویت، عملگرهای منطقی، عملگرهای بیت‌وایز
- d. آشنایی با دستورات شرطی، بازگشتی و تکرار
- e. داده ساختار ها مانند List, Tuples, Dictionary, Set
- f. آشنایی با نوشتن، اجرا و پیاده سازی تابع ها
- g. آشنایی با کتابخانه‌های پایتون

3. پیش پردازش و پاکسازی متن

- a. پیش پردازش متن چیست؟
- b. پاکسازی متن
- c. Tokenization
- d. Stemming and Lemmatization
- e. Part-of-Speech Tagging (POS Tagging)
- f. Named Entity Recognition

4. الگوریتمها و تکنیکها

- a. Bag of Words - Term Frequency
- b. TF/IDF
- c. word2vec
- d. مدل سازی موضوعات

f. خوشه بندی

i. vectorizing

ii. K-means

iii. مصورسازی

g. دسته بندی

i. vectorizing

ii. Naïve bayes

iii. ارزیابی نتایج با ماتریس آشفستگی

5. انجام پروژه

شیوه ارائه آموزش

روشها ، تکنیک ها و یا مدلهایی که برای تدریس مباحث درسی در این واحد قرار است از آنها استفاده شود.

سخنرانی تعاملی همراه با ارایه اسلاید

حل مساله

کار عملی

پرسش و پاسخ

شیوه ارزیابی دانشجو

نحوه ارزیابی در طول ترم و آزمون پایان ترم لازم است با ذکر نوع آزمون و همینطور ذکر بارم نمره به تفکیک هر یک از بخش های ارزیابی اعم از نمره حضور غیاب ، انجام تکلیف، مشارکت در بحثهای کلاسی، انجام کار گروهی و کوئیز ها و ... لازم در این بخش بطور دقیق و شفاف مشخص شود.

- حضور منظم و مشارکت در بحثهای کلاسی

- انجام تکالیف کلاسی در طول ترم

- ارائه پروژه نهایی

حداقل نمره قبولی برای این درس : بر اساس کوریکولوم

تعداد ساعات مجاز غیبت برای این واحد درسی : طبق آئین نامه های آموزشی مصوب

منابع آموزشی

منابعی که قرار است سوالات آزمون از آنها طرح شوند باید لیست شوند. اگر قرار است برخی منابع برای مطالعه بیشتر دانشجویان به آنها معرفی شوند ، لیست آنها باید بطور جداگانه در ادامه لیست شود.

- a. Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining. ChengXiang Zhai
- b. python software

c. تمامی کتب و راهنماهای آموزش پایتون

منابع آموزشی برای مطالعه بیشتر

از این منابع نمی توان سوال آزمون طرح کرد. این منابع صرفا به منظور تعمیق یادگیری دانشجویان به آنها معرفی می شوند.

فرصت های یادگیری

معرفی فرصتهای احتمالی و فراهم شده توسط اعضای هیئت علمی گروه / دانشگاه/ سایر دانشگاهها در طول ترم برای یادگیری بیشتر شامل لیست کارگاهها، وبینارها، کنفرانس ها ، ژورنال کلاب ها و به فراگیران

- شرکت فعال در کارگاه ها وبینارهای مرتبط با حوزه برنامه نویسی پایتون، داده کاوی و متن کاوی

اطلاعات تماس

مدرس / مدرسین دوره (تلفن ، ایمیل و):

- دکتر شفیع حبیبی
- Email: shafihabibi@gmail.com
- شماره تماس: ۰۴۱۳۱۷۷۵۹۶۳
- دکتر عذرا دائی
- Email: azradaei@gmail.com
- شماره تماس: ۰۴۱۳۱۷۷۵۹۶۴