

به نام خداوند بخشنده مهربان



تمرین و پروژه شماره ۲

درس داده کاوی

عنوان: تمرین و پروژه‌های تحویلی درس داده کاوی

استاد درس: حاجی محمدی

تاریخ تحویل: ۱۴۰۱/۳/۲۰

## تمرین های تحویلی

- I. تمرین های کتاب فصل ۳ شماره های ۱، ۳، ۴، ۱۰، ۱۶؛
- II. تمرین های کتاب فصل ۴ شماره های ۲، ۶، ۷، ۱۴؛
- III. تمرین های کتاب فصل ۵ شماره های ۲، ۵، ۷، ۸، ۹؛
- IV. تمرین های کتاب فصل ۶ شماره های ۱، ۴، ۵، ۱۳، ۱۴، ۲۲، ۲۳؛
- V. تمرین های کتاب فصل ۸ شماره های ۲، ۴، ۷، ۱۲؛
- VI. تمرین های کتاب فصل ۹ شماره های ۲، ۵، ۱۶، ۱۸؛

## پروژه های تحویلی

- I. استفاده از شبکه عصبی برای دسته بندی تصاویر  
هف از این تمرین تشخیص افرادی است که عینک به چشم دارند. شبکه عصبی باید بتواند با دریافت یک تصویر تشخیص دهد که آیا آن شخص عینک به چشم دارد یا نه؟

### سوالات

- 1) مجموعه داده را به مثالهای آموزشی و مجموعه تست تقسیم کنید. نحوه توزیع مثالها و نسبت هر دسته را مشخص کنید.
  - 2) یک شبکه دو یا سه لایه را با روش Backpropagation آموزش دهید. دلایل خود برای انتخاب ساختار شبکه وسایر پارامترهای طراحی را ذکر کنید.
  - 3) دقت شبکه را برای مجموعه داده آموزشی و مجموعه تائید بدست آورید
  - 4) آموزش شبکه را برای شبکه هائی با تعداد 2و4و10و50 گره پنهان تکرار کرده و با استدلال در مورد نتایج بهترین شبکه را از لحاظ دقت و زمان آموزش مشخص کنید
  - 5) اثر افزایش لایه های شبکه را بررسی کرده و نتیجه را گزارش کنید.
- مجموعه تصاویر پایگاه داده آموزشی را به صورت یک فایل فشرده به نام faces\_4 ارسال می شود.
- II. کد مربوط به یادگیری نمونه در بسته نرم افزاری WEKA را مطالعه نموده و الگوریتم های زیر را بررسی نمایید.

IB1: nearest neighbor classification

IBk: k-nearest neighbors' classification

## LWL (locally weighted learning)

به دلخواه خود یک مجموعه آموزشی در نظر بگیرید و الگوریتم‌های بالا را مورد ارزیابی قرار دهید.

*آنچه که تحویل می‌دهید:*

1. گزارشی به فرمت فایل PDF شامل پاسخهای داده شده هر یک از سوالات به همراه فایلها و سورس کدهای لازم
2. فایل pdf و کد را بصورت یکجا در قالب یک فایل zip در سامانه کورسز آپلود کنید (نام فایل = شماره دانشجویی & شماره دانشجویی HW2#1).

دانشجویان عزیز توجه داشته باشند که برای بوجود آمدن امکان ارزیابی صحیح کار انجام شده، لازم است تا هر دانشجو آزمایشات را جداگانه و یا فقط در گروه‌های تعیین شده انجام داده و نتایج جداگانه ای را نیز ارائه نماید.

**توجه:** هرگونه شباهت در گزارش و پاسخ تشریحی به منزله تقلب است و کل نمره تمرین صفر است. (میتوانید از اینترنت به عنوان منبع کمکی هم در سوالات تشریحی و هم در سوالات پیاده سازی استفاده کنید با ذکر منبع و ارجاع به آن، اما کپی برداری ممنوع است و نمره ی صفر تعلق میگیرد.) به گزارشات یکسان یا مشابه، نمره ای تعلق نمی گیرد.

در ضمن بصورت تصادفی از یک و یا چند دانشجو خواسته خواهد شد تا کار انجام شده را در کلاس درس توضیح بدهند.

امکان تمدید زمان ارسال پروژه وجود ندارد. لذا لطفا دوستان قبل از 59 : 23 تاریخهای مقرر، نسبت به بارگذاری پروژه اقدام فرمایند.

هر گونه سوال در مورد نحوه استفاده از نرم افزار Weka ، پایتون و یا سوالات را با تدریس یاران درس در میان بگذارید.

**موفق باشید**