

بنام خدا

تسک اول دوره هوش مصنوعی

آشنایی با الگوریتم‌های با نظارت (رگرسیون خطی، درخت تصمیم، جنگل تصادفی)

بخش اول:

با توجه به منابع ذکر شده در [1]، و دیگر منابعی که در اختیار دارید. الگوریتم‌های یادگیری ماشین رگرسیون خطی^۱، درخت تصمیم^۲ و جنگل تصادفی^۳ را بصورت کامل و خلاصه شرح دهید. روند انجام دسته‌بندی^۴ صورت گرفته توسط این مدل‌ها را شرح داده و داده‌هایی که این مدل‌ها برای آن‌ها مناسب هستند را توصیف کنید.

بخش دوم:

با استفاده از منابع موجود در [2] و دیگر منابع در دسترس، داده‌های مناسبی برای پیاده‌سازی الگوریتم‌های ذکر شده در بخش ۱ پیدا کنید. سپس، این الگوریتم‌ها را بر روی داده‌های یافت شده پیاده کنید و نتیجه‌ی حاصل از تست الگوریتم را بعنوان پاسخ برگردانید. آیا از یک پایگاه داده می‌توان برای هر سه الگوریتم استفاده کرد؟ چرا؟ الگوریتم‌هایی که می‌توانند روی یک پایگاه داده آموزش ببینند را با یکدیگر مقایسه کنید (از نظر دقت روی همان پایگاه داده و سرعت آموزش و تست).

پیوست:

برای تحویل تسک‌ها، در ابتدا یک پوشه در IUTBox ایجاد کنید. نام پوشه بصورت زیر باشد.

FirstName_LastName_StudentID_AI

این پوشه را با من با آدرس ایمیل a.iranidoust@ec.iut.ac.ir در IUTBox به اشتراک بگذارید. (اگر به مشکل برخوردید، پوشه را با لینک دسترسی به اشتراک بگذارید و لینک را برای من به آدرس ذکر شده ارسال کنید).

داخل پوشه‌ی ایجاد شده، یک پوشه بنام پیاده‌سازی‌ها و یک پوشه بنام گزارش‌ها ایجاد کنید. داده‌ها و کدهای پیاده‌سازی شده را فشرده کرده و داخل پوشه‌ی پیاده‌سازی‌ها قرار بدید. گزارش‌های نوشته شده را داخل پوشه‌ی گزارش‌ها قرار بدید.

¹ Linear Regression

² Decision Tree

³ Random Forest

⁴ Classification

[1]: <https://iutbox.iut.ac.ir/index.php/s/W8icntGFefbnqDT>

[2]: <https://www.kaggle.com/datasets>