

به نام خدا



دانشگاه علم و صنعت ایران

دانشکده مهندسی کامپیوتر

گروه هوش مصنوعی و رباتیک

پروژه درس شناسایی الگو

استاد درس:

دکتر مرتضی آنالویی

پروژه دوم

تدریس یار:

سحر سرکار

نیم سال اول تحصیلی ۱۴۰۱-۱۴۰۲

## مقدمه

هدف از این پروژه آشنایی با مفاهیم اولیه‌ی رده‌بندی است. در این پروژه از شما می‌خواهیم تا الگوریتم‌های رده‌بندی را بر روی مجموعه داده‌های واقعی پزشکی پیاده‌سازی کرده تا علاوه بر آشنایی و تسلط بر الگوریتم‌ها، با چالش‌های موجود در مجموعه داده‌های واقعی نیز آشنا شوید.

## مجموعه دادگان

دو فایل csv به‌عنوان مجموعه دادگان این پروژه به پیوست ارسال شده است. مجموعه داده‌ی Lateralities شامل اطلاعاتی برای تشخیص بیماری صرع است که از یکی از بیمارستان‌های تهران دریافت شده و مجموعه داده‌ی COVID\_Dataset نیز از تصاویر CT بیماران دارای Covid-19 بیمارستان غیائی تهران به دست آمده است.

## شرح پروژه

این پروژه از دو بخش تشکیل می‌شود که در ادامه به توضیح هر یک از این بخش‌ها می‌پردازیم.

### بخش اول

در این بخش می‌خواهیم از مجموعه دادگان تشخیص بیماری صرع که از یکی از بیمارستان‌های تهران دریافت شده است استفاده کنیم. این مجموعه داده بر اساس MRI‌های دریافتی از بیماران تهیه شده است. از شما می‌خواهیم تا با استفاده از تکنیک Cross Validation با مقادیر  $K = 5, 7, 10$ ، دقت<sup>۱</sup> حاصل از پیاده‌سازی الگوریتم AdaBoost بر روی این مجموعه داده را مقایسه کنید. برای پیاده‌سازی نکات زیر در نظر گرفته شود.

- در مجموعه داده، L به معنای چپ، R به معنای راست، UL به معنای تمایل به چپ، UR به معنای تمایل به راست و U را میانی در نظر می‌گیریم. ابتدا به هر یک کلاسی تخصیص دهید (می‌توانید L و UL را دو کلاس جدا و یا دو کلاس یکسان در نظر بگیرید).
- برچسب داده‌ها در ستون Overall\_Lateralities\_NO آورده شده است.

### بخش دوم

در این بخش می‌خواهیم از مجموعه دادگان استخراج شده از تصاویر CT بیماران دارای Covid-19 برای افتراق نواحی دارای عفونت و نواحی غیرعفونی استفاده کنیم. الگوریتم‌های دسته‌بندی AdaBoost و Decision Tree را بر روی این مجموعه داده‌ی برچسب‌گذاری شده پیاده‌سازی کرده و نتایج حاصل از این دو الگوریتم را مقایسه کنید.

---

<sup>1</sup> Accuracy

## نکات تکمیلی

- انجام پروژه به صورت گروهی مجاز نمی باشد.
- گزارش پروژه می بایست شامل توضیح تمام مراحل پیاده سازی و گزارش نتایج حاصل باشد.
- سوالات خود را می توانید در گروه تلگرامی درس از کمک مدرس مربوطه بپرسید و در صورت عدم دسترسی به تلگرام می توانید از طریق ایمیل زیر مطرح کنید.

[sarkarsahar@gmail.com](mailto:sarkarsahar@gmail.com)

- پاسخ خود را در یک فایل فشرده که شامل یک نوت بوک (.ipynb) و فایل گزارش پروژه (PDF) است، به فرمت زیر در سامانه ال ام اس ارسال کنید.

PJ2\_YourStudentID\_YourName.zip