

به نام خدا

مسابقه شماره 2 دوره هوش مصنوعی فناوری

تابستان 1400

در مباحث منابع انسانی یکی از مهم ترین فاکتور های موفقیت یک شرکت نرخ ترک کارکنان (Attrition Rate) است. از عوامل ترک شرکت توسط کارکنان میتوان به دلایل شخصی، انگیزه های حرفه ای، عدم تطابق با شغل و مشکلات در محیط کار اشاره کرد.

یکی از اهداف مهم واحد های منابع انسانی کم کردن این نرخ است. در این مسابقه سعی میکنیم مدلی طراحی کنیم تا بتواند احتمال ترک هر کارمند را با توجه به ویژگی های مختلف مانند واحد شغلی، حقوق و ... را پیشبینی کند.

در این مسابقه هم مانند مسابقه قبلی دو دیتاست در اختیار شما قرار میگیرد که اولی به نام hr\_train.csv برای آموزش دادن مدل مناسب به کار میرود و دیتاست hr\_test.csv برای ارزیابی عملکرد شما.

در این دیتاست ستون Attrition به عنوان label و یا target مورد استفاده قرار میگیرد و سایر ستون ها به عنوان feature .

#### معیار ارزیابی:

برای ارزیابی الگوریتم ها ما از معیار f1 score با رابطه زیر استفاده میکنیم:

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

**نکته شماره 1:** شما مختارید برای رسیدن به بهترین الگوریتم از تمامی روش هایی که در کالس ذکر شده و یا نشده است استفاده کنید. بنابراین توصیه میشود با جستجو در اینترنت روش های متفاوت را یاد بگیرید و در کار خود بکار بگیرید.

**نکته شماره 2:** این دیتاست به طور عمومی در وبسایت کگل به آدرس <https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset> موجود است و میتوانید با مطالعه گفتگو ها و پیاده سازی سایر افراد ایده های مناسبی بگیرید.

**نکته شماره 3:** حتما از کپی کردن از کار افراد خودداری کنید و اگر از پیاده سازی شخص دیگری استفاده کردید حتما منبع آن را ذکر کنید و سعی کنید ایده جدیدی به آن بافزایید.

**نکته شماره 4:** از آنجایی که کل این دیتاست در اینترنت موجود است، شخص برنده موظف است که کد هایی که منجر به نتیجه نهایی میشود را پس از پایان مسابقه ارسال کند و برای دریافت جایزه این نتایج باید فقط با استفاده از دیتاست ترین حاصل شده باشند.

**نکته شماره 5:** به برنده این مسابقه در پایان زمان برگزاری جایزه ای تقدیم خواهد شد.

**نکته شماره 6:** برنده مسابقه در اولین جلسه پس از مسابقه ارائه کوتاهی از روش های به کار برده شده ارائه خواهد کرد.

**نکته شماره 7:** یک جایزه اضافی برای یک نفر که یک کار تحلیلی جالب روی این داده ها انجام دهد نیز در نظر گرفته شده است. این کار تحلیلی میتواند ارائه visualization های جذاب و تحلیل داده ها برای پیدا کردن ارتباط بین feature های مختلف باشد.

**نکته شماره 8:** شما مختارید در طول برگزاری مسابقه سوالات خود را از همکالسی ها یا مدرس کالس بپرسید. (توصیه میشود این سوالات در گروه عمومی اسلک مطرح شود تا هم دیگر همکاسی ها از بحث ها استفاده کنند و هم امتیاز شما برای مشارکت عمومی در کلاس بالا رود).