

«به نام خدا»

عنوان : گزارش کار پروژه‌ی یادگیری ماشین، استعداد یابی ورزش کشتی

انجام دهنده : امیر مهدی صالح

شماره دانشجویی : ۴۰۲۱۲۳۴۰۱۱۹۰۳۹

مقدمه

در سال‌های اخیر، استفاده از هوش مصنوعی و یادگیری ماشین در تحلیل داده‌های ورزشی رشد چشمگیری داشته است. یکی از کاربردهای مهم این حوزه، استعدادیابی ورزشی و پیش‌بینی نتایج مسابقات می‌باشد. هدف از این پروژه، طراحی یک سیستم هوشمند است که بتواند با دریافت داده‌های متنه مربوط به ویژگی‌های ورزشکاران، رشته ورزشی مناسب (در این پروژه کشتی) را پیشنهاد دهد و در مراحل بعدی امکان توسعه به داده‌های تصویری و ویدئویی (Pose Detection) را داشته باشد.

در این پروژه ابتدا از داده‌های متنه استفاده شده و یک مدل یادگیری ماشین آموزش داده شده است که بر اساس ویژگی‌های فیزیکی و آماری ورزشکار، خروجی مناسب را پیش‌بینی می‌کند. در نهایت یک رابط کاربری ساده برای استفاده از مدل طراحی شده است.

2. معرفی دیتاست

برای این پروژه یک دیتاست ایرانی از اطلاعات ورزشکاران کشتی تهیه شد که شامل نمونه‌هایی از ویژگی‌های فیزیکی و آماری ورزشکاران می‌باشد. این دیتاست به صورت فایل CSV ذخیره شده شامل ستون‌هایی مانند:

سن

قد

وزن

سابقه تمرین

تعداد برد

تعداد باخت

سطح آمادگی جسمانی

برچسب خروجی (موفق یا ناموفق / سطح مهارت / نتیجه مسابقه)

هدف از استفاده از دیتاست ایرانی، افزایش دقیقی مدل برای جامعه هدف و دریافت امتیاز بیشتر طبق دستور پروژه بوده است.

3. پیش‌پردازش داده‌ها

3-1. پاکسازی داده‌ها

در مرحله اول داده‌ها از نظر وجود مقادیر گمشده (NaN) بررسی شدند. در صورت وجود مقادیر ناقص، این داده‌ها یا حذف شدند یا با مقادیر میانگین جایگزین شدند.

3-2. انتخاب ویژگی‌ها

ستون‌های غیرضروری حذف شدند و فقط ویژگی‌هایی که در پیش‌بینی موثر بودند نگه داشته شدند.

3-3. نرمال‌سازی داده‌ها

برای جلوگیری از اثر مقیاس متفاوت ویژگی‌ها (مثلاً قد و وزن)، داده‌ها با استفاده از روش استانداردسازی (StandardScaler) نرمال‌سازی شدند.

3-4. تبدیل برچسب‌ها

چون خروجی مدل به صورت متئی بود (مثلاً سطح مهارت)، با استفاده از LabelEncoder این مقادیر به اعداد تبدیل شدند تا برای مدل قابل فهم باشند.

4. تقسیم داده‌ها

داده‌ها به دو بخش تقسیم شدند:

80% برای آموزش (Training)

20% برای آزمون (Testing)

این کار با استفاده از تابع train_test_split انجام شد تا عملکرد مدل به صورت واقعی ارزیابی شود.

5. انتخاب مدل یادگیری ماشین

در این پروژه از الگوریتم Random Forest Classifier استفاده شده است. دلیل انتخاب این مدل: دقت بالا

مقاومت در برابر نویز

توانایی کار با داده‌های عددی

کاهش مشکل overfitting نسبت به درخت تصمیم ساده

6. آموزش مدل

مدل با استفاده از داده‌های آموزش آموزش داده شد. سپس با داده‌های آزمون تست شد و مقدار دقت (Accuracy) محاسبه گردید:

Accuracy ≈ 0.88

این مقدار نشان می‌دهد که مدل توانسته است حدود 88٪ داده‌ها را به درستی پیش‌بینی کند که برای یک پروژه دانشگاهی مقدار مناسبی است.

7. ذخیره مدل

پس از آموزش، مدل و LabelEncoder با استفاده از کتابخانه joblib ذخیره شدند:

text_model.pkl

label_encoder.pkl

این کار باعث می‌شود بدون نیاز به آموزش مجدد، مدل در برنامه نهایی استفاده شود.

8. طراحی رابط کاربری (UI)

برای استفاده آسان از مدل، یک رابط کاربری تحت وب با استفاده از Streamlit طراحی شد. این رابط شامل:

فرم ورود اطلاعات ورزشکار

دکمه پیش‌بینی

نمایش نتیجه پیش‌بینی به کاربر

کاربر با وارد کردن ویژگی‌ها می‌تواند نتیجه پیش‌بینی را مشاهده کند.

9. اجرای برنامه در محیط ابری

از Google Colab برای اجرای برنامه استفاده شد.

به منظور در دسترس قرار دادن برنامه برای ارائه، از Cloudflare Tunnel استفاده گردید تا یک لینک عمومی ایجاد شود.

این لینک امکان مشاهده و تست برنامه را بدون نصب نرم‌افزار فراهم می‌کند.

10. Pose Detection (توضیح مرحله دوم پروژه)

در فاز دوم پروژه، امکان توسعه به داده‌های تصویری و ویدئویی وجود دارد. در این مرحله می‌توان از مدل‌هایی مانند:

MediaPipe

OpenPose

برای استخراج نقاط کلیدی بدن (Keypoints) استفاده کرد و این داده‌ها را به مدل یادگیری ماشین داد تا استعدادیابی دقیق‌تر انجام شود.

11. نتایج پروژه

طراحی یک سیستم استعدادیابی برای ورزش کشتی

دقت مدل حدود ۸۸٪

طراحی رابط کاربری

امکان توسعه به داده‌های تصویری

استفاده از دیتاست ایرانی

حجم پایین پروژه

سرعت مناسب اجرا

12. محدودیت‌ها

دیتاست نسبتاً کوچک

استفاده از داده‌های متنی به جای داده‌های تصویری

عدم استفاده از شبکه‌های عصبی عمیق به دلیل محدودیت سخت‌افزاری

13. پیشنهاد برای کارهای آینده

افزودن تحلیل ویدئویی مسابقات

استفاده از CNN برای تصویر

افزایش حجم دیتاست

استفاده از داده‌های بیومتریک

افزودن پیش‌بینی نتیجه مسابقه بین دو کشتی‌گیر

14. نتیجه‌گیری

در این پژوهه یک سیستم هوشمند برای استعدادیابی ورزشی در رشته کشتی طراحی شد که با استفاده از یادگیری ماشین و داده‌های متنی می‌تواند سطح مهارت و نتیجه مسابقه را پیش‌بینی کند. طراحی رابط کاربری باعث شده استفاده از سیستم برای کاربران ساده باشد. این پژوهه نمونه‌ای کاربردی از استفاده یادگیری ماشین در حوزه ورزش می‌باشد.