



دانشگاه شهید بهشتی  
Shahid Beheshti University

پروژه پایانی

درس یادگیری ماشین

مدرس: دکتر احمدعلی آبین



## استخراج سیکل رانندگی شهری تهران با استفاده از مدل مارکوف پنهان سلسله‌مراتبی (HHMM)

### تعریف پروژه

در این پروژه با بهره‌گیری از داده‌های ثبت‌شده رانندگی در شرایط واقعی شهری تهران، یک سیکل رانندگی بومی و استاندارد برای شهر تهران طراحی خواهد شد. داده‌ها شامل ۸۰ نمونه مسیر رانندگی است که توسط ۸۰ راننده مختلف، با سبک‌های متنوع (از رانندگی کاملاً اقتصادی تا بسیار پرمصرف) با یک خودروی دنده‌دستی و در مسیری ثابت به طول ۱۰ کیلومتر در تهران طی شده است. داده‌ها با فرکانس ۸۰۰ میلی‌ثانیه جمع‌آوری شده‌اند و شامل پارامترهایی مانند سرعت، دور موتور، دنده، درصد پدال گاز، موقعیت جغرافیایی، مصرف سوخت، شیب مسیر و متغیرهای مرتبط با عملکرد پیش‌ران هستند.

سیکل رانندگی (Driving Cycle) یک پروفایل سرعت-زمان است که رفتار رانندگی را در یک بازه زمانی مشخص بازنمایی می‌کند. این پروفایل معمولاً شامل توقف‌ها، شتاب‌گیری‌ها، رانندگی با سرعت ثابت، و کاهش سرعت است. سیکل رانندگی به عنوان ورودی تست‌های مصرف سوخت و آلایندگی خودرو در آزمایشگاه‌ها استفاده می‌شود.

در سطح جهانی، سیکل‌های استاندارد مانند WLTC (سیکل جهانی هماهنگ‌شده)، NEDC (اروپا)، JC08 (ژاپن) و FTP75 (آمریکا) برای تست مصرف سوخت و آلایندگی خودروها مورد استفاده قرار می‌گیرند. این سیکل‌ها ساختارهای زمانی خاصی دارند و شامل سه فاز اصلی هستند:

- فاز توقف (Idle)
- فاز رانندگی درون‌شهری (Urban)
- فاز رانندگی برون‌شهری (Extra-Urban)

با این حال، تاکنون برای ایران و به‌ویژه کلان‌شهر تهران، سیکل رانندگی بومی‌سازی‌شده‌ای که رفتار واقعی رانندگان را منعکس کند، توسعه نیافته است. هدف این پروژه طراحی چنین سیکلی، مبتنی بر داده‌های واقعی، با ساختار سه‌فازی و قابل مقایسه با سیکل‌های بین‌المللی است.



## روش کار

### ۱. پیش‌پردازش و برجسب‌گذاری اولیه:

- محاسبه ویژگی‌هایی مانند شتاب، توقف، تغییرات دنده، نرخ تغییر سرعت، زاویه شیب و مصرف لحظه‌ای سوخت.
- تعیین فازهای رانندگی با استفاده از قواعد ساده مبتنی بر سرعت و توقف طبق سیکل‌های استاندارد.

### ۲. مدل‌سازی با HMM و HHMM:

- آموزش مدل مارکوف پنهان معمولی (HMM) برای شناسایی ساختار توالی فازها.
- توسعه مدل مارکوف پنهان سلسله‌مراتبی (Hierarchical HMM) برای بازنمایی بهتر گذارهای بین فازهای اصلی و زیرالگوهای رفتاری
- تولید یک پروفایل زمانی رانندگی با ساختار استاندارد که شامل فازهای مختلف با درصد مشابه سیکل‌های قبلی باشد.
- هموارسازی پروفایل و ایجاد سیکل سرعت نهایی و رسم پروفیل سرعت و ذخیره در فایل CSV.

### ۳. تخمین دنده با HMM:

- آموزش مدل HMM برای تخمین دنده خودرو
- بازسازی توالی منطقی دنده در سیکل خروجی مرحله قبل، مناسب برای خودروهای دنده‌دستی.



## ورودی و خروجی مورد نظر برای انجام پروژه

### ورودی:

ورودی‌ها باید از داده‌های خام استخراج شوند و رفتار رانندگی را به خوبی نمایش دهند. در این پروژه، انتخاب مناسب متغیرها اهمیت زیادی دارد، چون HHMM باید بر اساس این ورودی‌ها فازهای رانندگی و الگوهای رفتاری زیرمجموعه را تشخیص دهد. بر اساس ویژگی‌های زیر می‌توانید ترکیبی از ورودی‌ها برای ساخت مدل HHMM خود انتخاب کنید. استفاده از حداقل ۳ ویژگی اجباری می‌باشد.

### ۱. دینامیک حرکت خودرو

- Speed : سرعت خودرو
- Acceleration : شتاب خودرو (متر بر مجذور ثانیه)
- Jerk : تغییرات شتاب

### ۲. کنترل راننده

- Throttle Position : موقعیت دریچه گاز
- Accelerator Pedal Position : موقعیت پدال گاز
- Brake Indicator : یا استنتاج از شتاب منفی

### ۳. وضعیت موتور

- Engine Speed (RPM) : دور موتور
- Current Gear Shift Position : دنده خودرو

### ۴. پارامترهای مصرف

- Fuel Consumption Rate : مصرف سوخت لحظه‌ای
- Trip Fuel Consumption : مصرف سوخت تجمعی

### ۵. شرایط محیطی

- Slope : شیب
- Altitude : ارتفاع از سطح دریا

### خروجی:

خروجی نهایی باید یک یا چند فایل سیکل یا سری زمانی شامل ستون‌های زیر باشد (فایل CSV) که نمودار سرعت-زمان و دنده-زمان نیز جداگانه رسم شده باشد.

- Time (s)
- Speed (km/h) (الزامی)
- Phase (Idle, Urban, Extra-Urban)
- Gear



## معیارهای ارزیابی و مقایسه با سیکل‌های جهانی

سیکل نهایی تهران از منظر فنی و آماری با سیکل‌های JC08، FTP-75، NEDC، WLTC مقایسه خواهد شد. هدف این است که ساختار و ویژگی‌های آماری سیکل پیشنهادی، حداکثر شباهت ممکن را با یکی از این سیکل‌های استاندارد (به‌ویژه از نظر طول سیکل، تعداد فازها و نسبت زمانی هر فاز) داشته باشد، در عین حال پارامترهای آن بر اساس داده‌های واقعی رانندگی در تهران تعیین شود.

- میانگین سرعت هر فاز و کلی (Mean Speed)
- تعداد توقف و مدت کل توقف (Idle Time %)
- نسبت زمانی هر فاز به کل سیکل (Phase Time Ratio)
- توزیع سرعت برای هر فاز و به صورت کلی (Speed Distribution Histogram)
- دامنه نوسانات سرعت و شتاب در هر فاز و کلی (Speed/Acceleration Variability)
- مسافت طی شده (KM) هر فاز و کل

برای تحلیل دقیق شباهت بین سیکل پیشنهادی تهران و سایر سیکل‌ها، از معیارهای زیر استفاده می‌شود:

### ۱. Kullback-Leibler Divergence (KLD)

برای سنجش میزان اختلاف آماری بین توزیع فازهای مختلف.

### ۲. Dynamic Time Warping (DTW)

برای تطبیق دینامیکی پروفایل‌های سرعت در طول زمان (معیار انطباق زمانی با امکان جابه‌جایی جزئی).

۳. سیکل تولید شده باید دارای تمام فازها باشد و از سرعت + شروع شده و به سرعت + ختم شود و همچنین باید فازهای تولید شده در یک تصویر مشخص شده باشد.

**توجه:** در انتهای نوت‌بوک یک سلول مجزا ایجاد کنید تا امکان اجرای موارد خواسته‌شده بر روی فایل‌های تست (که توسط گروه حل تمرین ارائه می‌شوند) فراهم گردد. نام ستون‌های موجود در فایل‌های تست مطابق با فایل‌های پیوست‌شده خواهد بود.



## نوآوری و اهمیت پروژه

- تولید سیکل رانندگی بومی شده برای تهران با استفاده از داده‌های واقعی و مدل‌های یادگیری آماری پیشرفته (HHMM)
- استفاده از روش‌های تخمین دنده مبتنی بر یادگیری ماشینی مناسب برای خودروهای دنده‌دستی
- بهره‌گیری از معیارهای فنی در کنار شاخص‌های آماری برای ارزیابی جامع شباهت سیکل
- کاربردپذیری در صنایع خودروسازی، سازمان ملی استاندارد، و پژوهش‌های حوزه انرژی و محیط‌زیست



## نکته مهم انجام پروژه :

در مدل‌های HMM، انتخاب متغیرهای مشاهده‌شده و حالات پنهان آزاد است و این انعطاف‌پذیری یکی از نقاط قوت مدل است. با وجود متغیرهای مختلف مثل سرعت، دور موتور و شیب جاده، استفاده صرف از یک متغیر مثل سرعت نمی‌تواند رفتار پیچیده رانندگان را به درستی مدل کند.

این پروژه برای درس یادگیری ماشین مقطع ارشد تعریف شده و انتظار می‌رود هر فرد با خلاقیت خود، متغیرها و روش‌های متفاوتی انتخاب کند. استفاده همه از متغیرهای یکسان یا مدل، تقلب محسوب شده و ارزش علمی کار را پایین می‌آورد.

بنابراین، خلاقیت در انتخاب ویژگی‌ها و طراحی مدل‌ها برای دستیابی به نتایج معتبر و کاربردی ضروری است.



## نکات حائز اهمیت در تحویل پروژه

- ۱- تحویل پروژه از طریق سامانه درس‌افزار می‌باشد.
- ۲- کد باید در قالب نوت‌بوک و نتایج آن قابل مشاهده باشد، عدم توجه به این موضوع باعث کسر نمره می‌شود.
- ۳- موارد خواسته شده یعنی فایل‌های CSV و تصاویر خروجی برای سرعت و دنده به همراه فایل نوت‌بوک اپلود را در قالب یک فایل فشرده بارگذاری کنید و نام فایل به صورت زیر باشد.

FML404-StudentName-StudentID

- ۴- پروژه به صورت تک نفره است و هرگونه تقلب از سایر افراد تقلب محسوب می‌شود. کمک از AI‌های مرسوم به شرط تسلط کافی مانعی ندارد.
- ۵- در صورت داشتن هرگونه سوال با ایدی‌های زیر در ارتباط باشید:

@sirmmp

@ali\_es779

پیروز باشید.