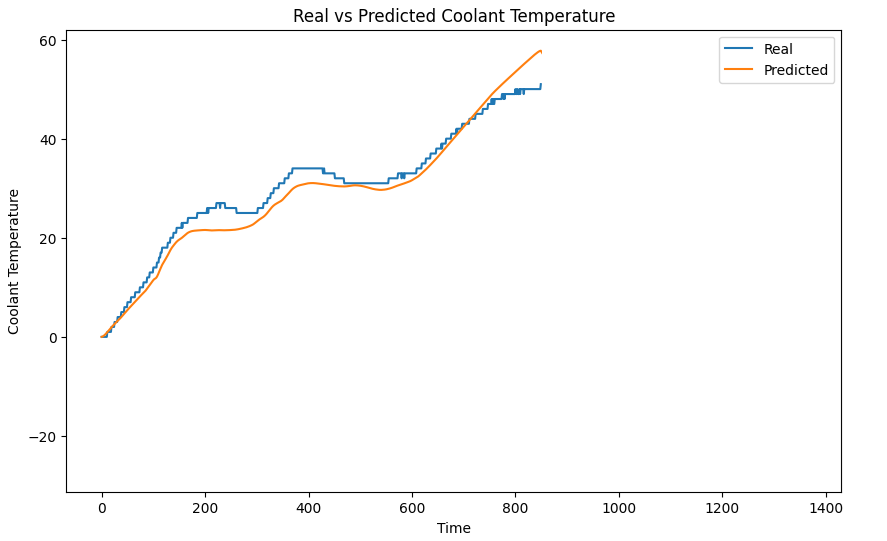
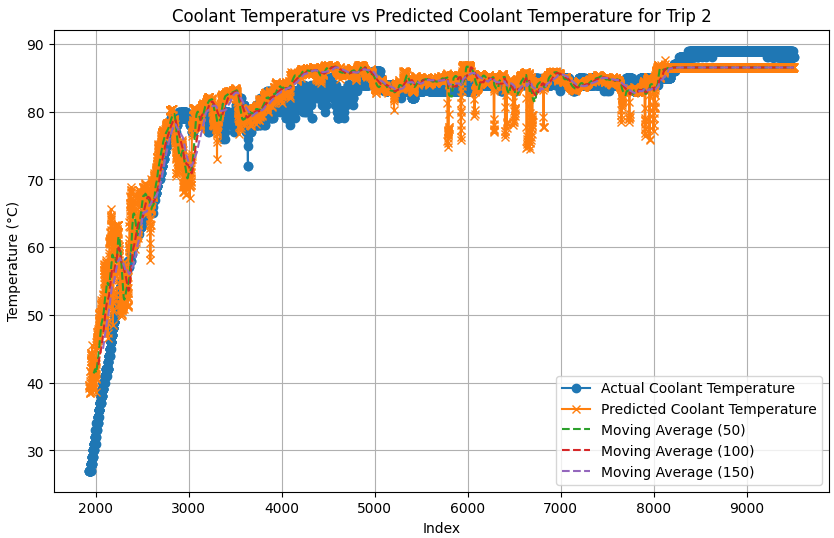
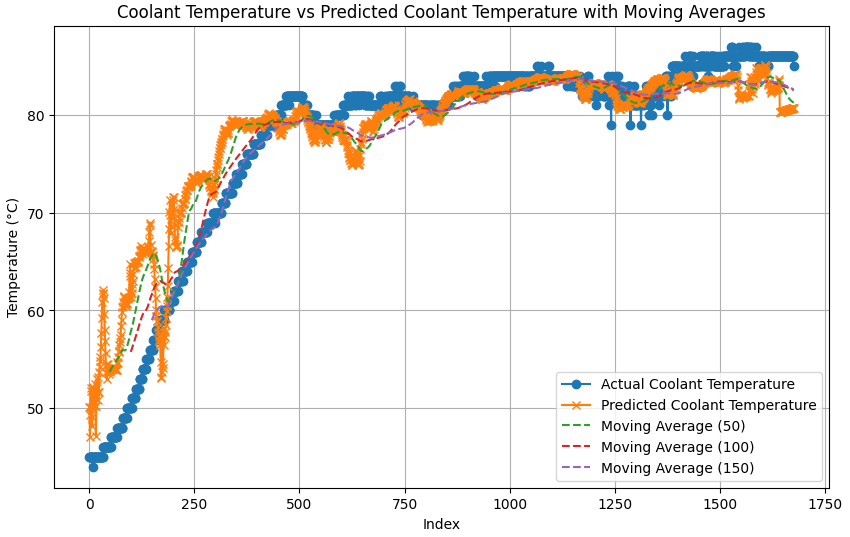
به نام خدا

در این مینی پروژه، هدف پیش بینی الگوی افزایش دما، تا رسیدن به حد ثابت می‌باشد، در گام‌های ابتدایی این پروژه از داده‌های خودرو 791 استفاده شد، سپس بر روی آن عملیات‌های مختلف داده افزایی انجام شد، سپس با داده‌های جدید که تعداد آن‌ها به اندازه کافی بود از شبکه BLSTM برای پیش‌بینی استفاده شد، در ادامه این شبکه بهبود یافت و توانایی ورودی گرفتن از فایل‌هایی با سایز داینامیک را بدست آورد(تا قبل آن باید همه تبدیل به یک بازه محدود 600تایی می‌شدند)، در ادامه خروجی پیش بینی مدل داینامیک را میبینیم .

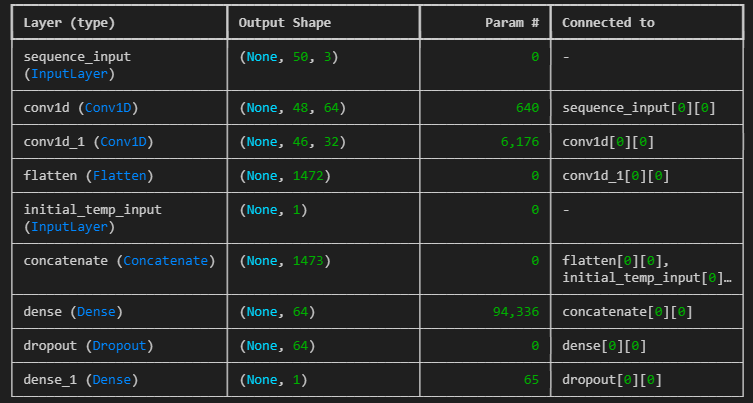
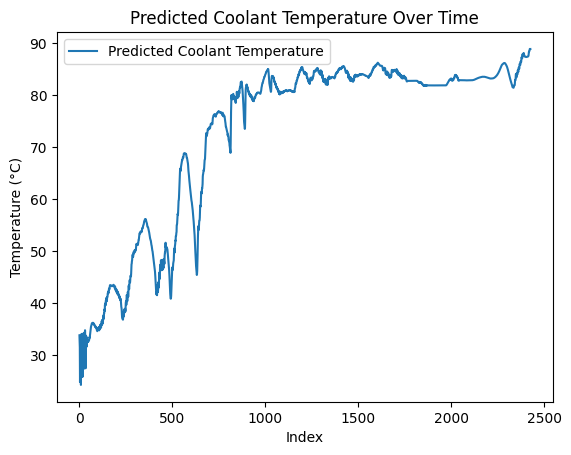


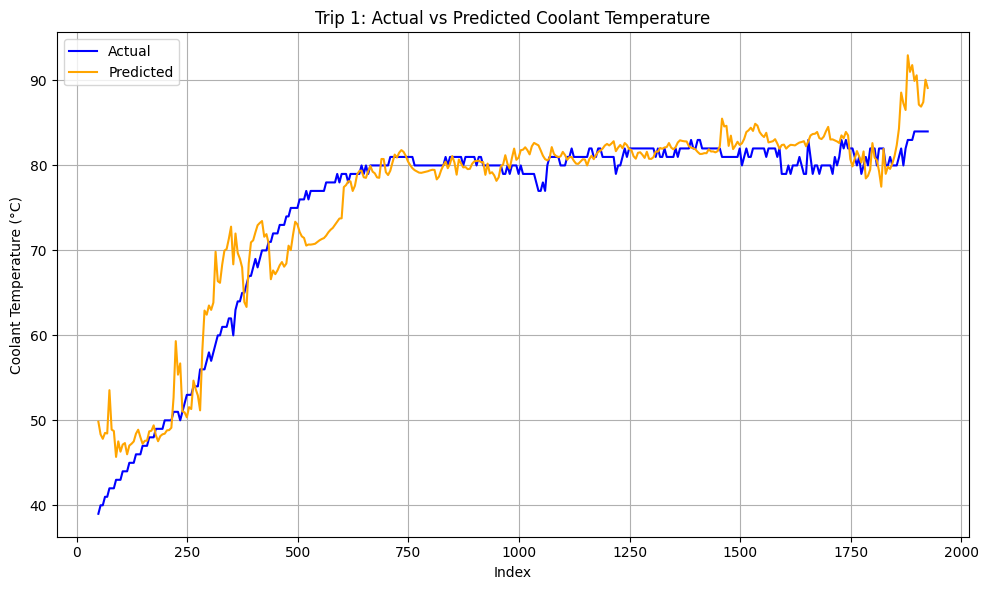
پیش بینی مدل داینامیک

در گام بعدی به سراغ الگوریتم‌های ساده‌تر و کم‌هزینه‌تر رفتیم، در گام ابتدایی برای دماهای استارت مختلف مدل‌های متفاوتی آموزش داده شد، سپس با توجه به پارامتر‌های تعریف شده این پیش‌بینی صورت گرفت که در ادامه نتایج آن را می‌بینیم.

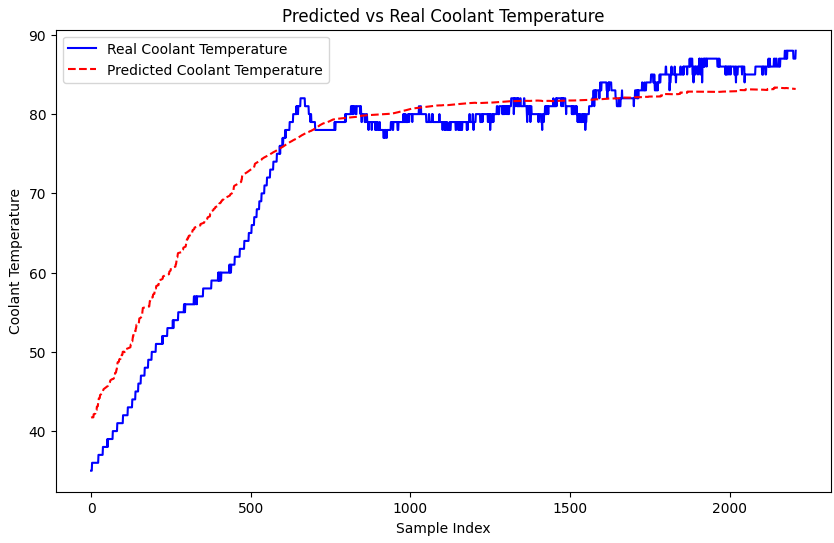
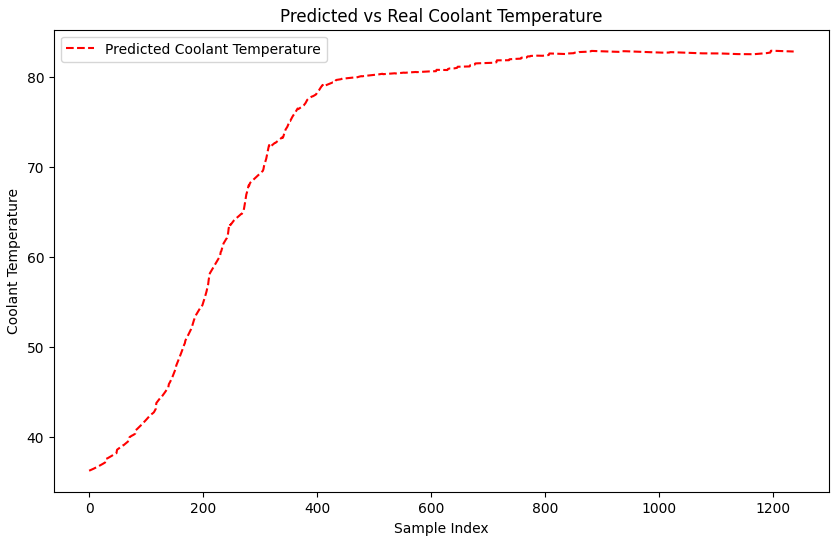


در ادامه به سراغ شبکه CNN یک بعدی رفتیم، ساختار و همچنین پیش بینی آن برای سیکل NEDC را در ادامه مشاهده می‌نمایید.

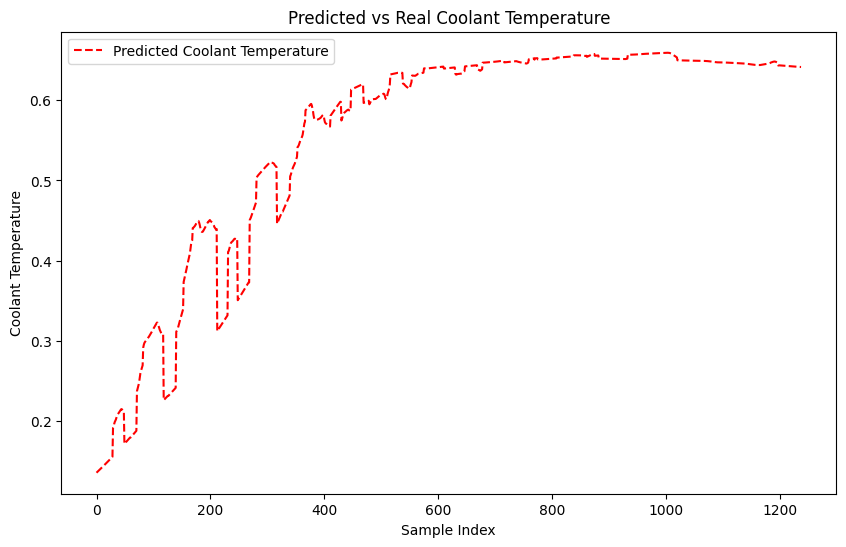




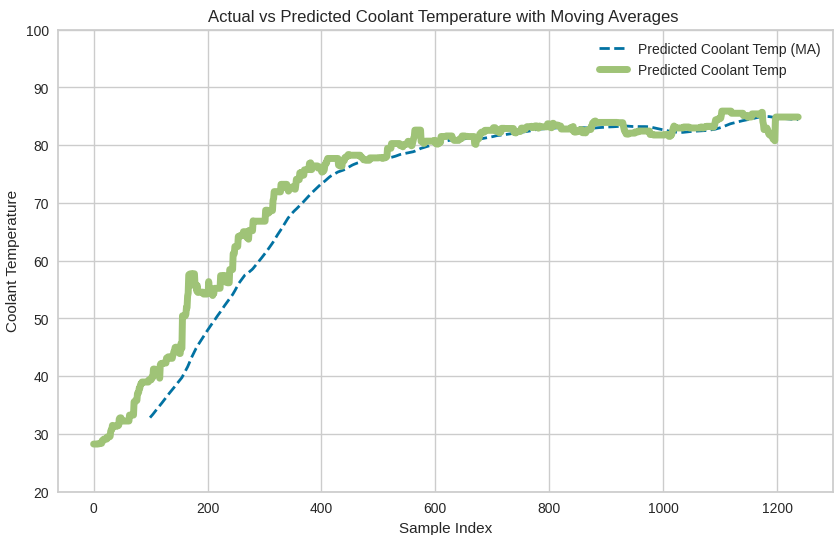
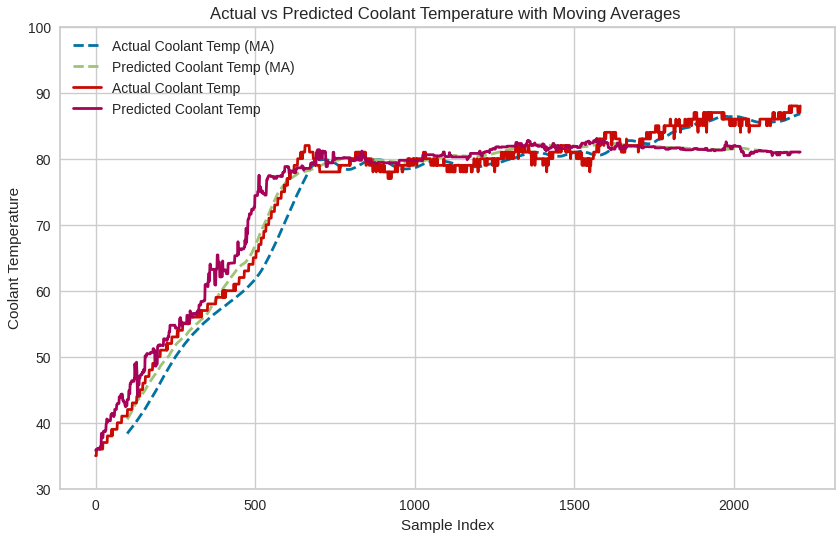
در ادامه یک ساختار دیگر در جهت بهبود توانایی یافتن ویژگی‌های محلی و کلی ایجاد شد، راه حل مدنظر برای این قسمت استفاده از دو پنجره جدید بود که خروجی بهبود یافته را در ادامه می‌بینیم:



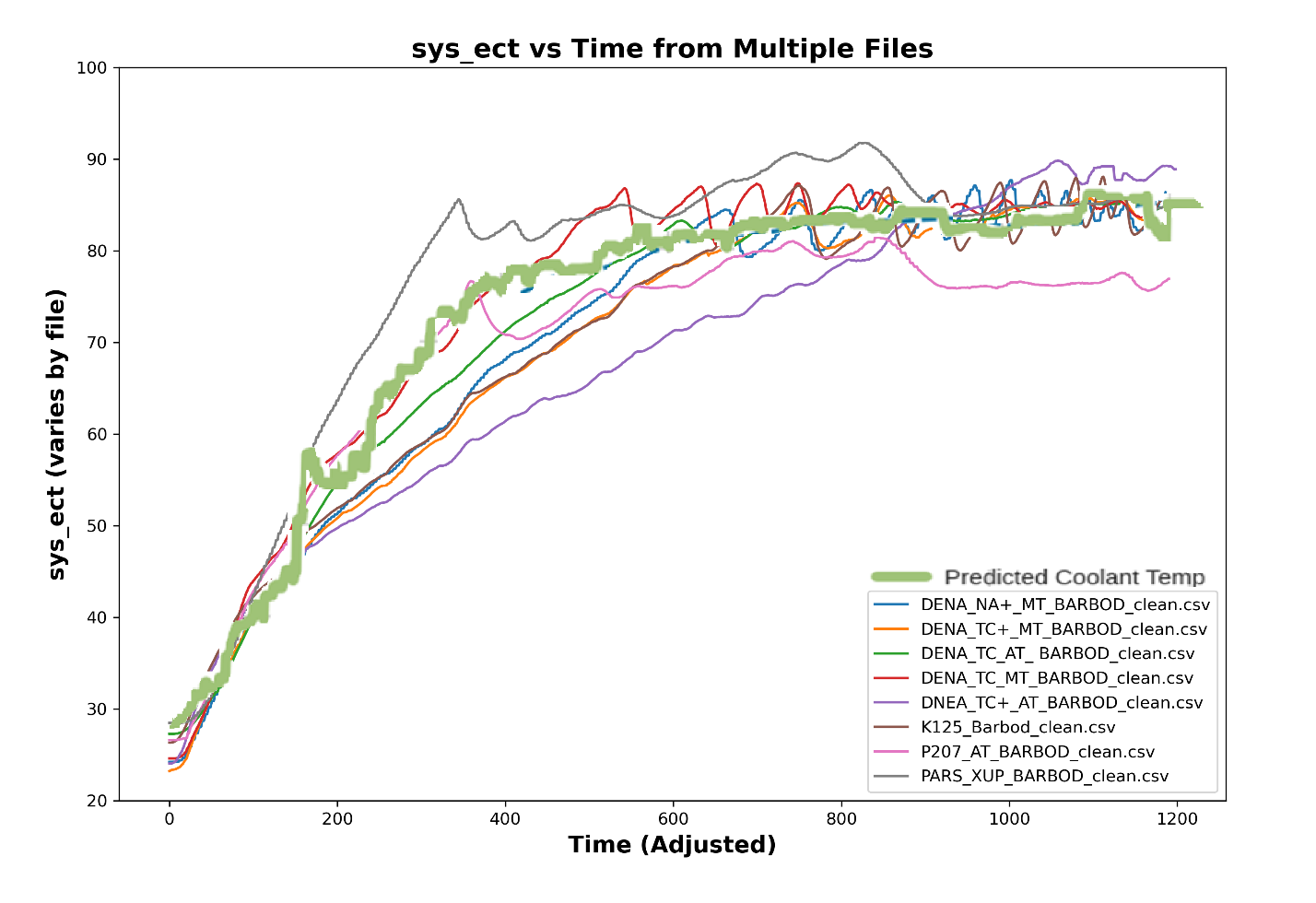
در نهایت به سراغ روشی با زمان آموزش کمتر رفتیم، با توجه به توانایی بالای شبکه XGBoost علی رغم نیاز به زمان یادگیری بالا، توانایی و قابلیت یادگیری بالایی دارد در ابتدا ویژگی‌های ورودی زمان، سرعت، دنده و دمای اولیه خودرو بودند، مشکل این مدل هم جنس نبودن داده‌ها بود چون زمان به صورت جمع شونده ولی سرعت و دنده به صورت لحظه‌ای بود از این رو تغییرات سرعت و دنده در پیشبینی دیده می‌شد. شکل زیر مثالی از این نمونه هست.



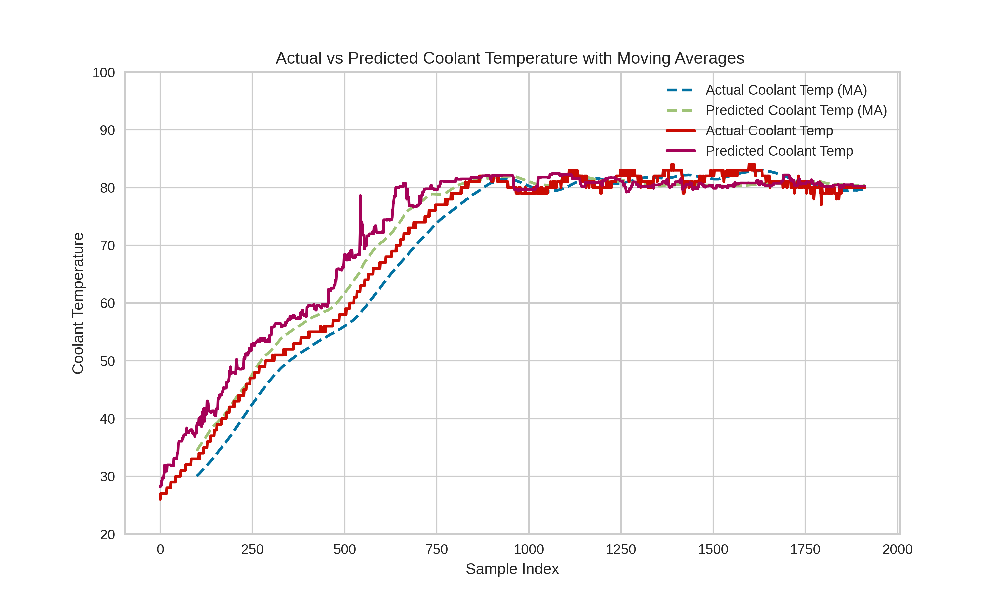
بعد از دیده شدن این مشکل باید تغییراتی در ورودی و ویژگی‌ها ایجاد می‌کدیم تا پیشبینی ما بهتر شود. یکی از ایده‌های مطرح شده استفاده از مجموع تغییرات سرعت و دنده نسبت به لحظه 0 بود یعنی اگر سرعت در لحظه 0 برابر0 در لحظه 1 برابر 10 و در لحظه 3 برابر 5 بود مجموع تغییرات برای لحظه 3 برابر 15 است. برای دنده هم به همین صورت عمل کردیم و مجموع تغییرات را در لحظه جایگزین کردیم. بعد از انجام این تغییرات دیگر تغییرات سرعت و دنده به صورت شکل بالا قابل مشاهده نبود و دقت مدل نیز بهتر شد. در ادامه نتایج را مشاهده می‌کنیم.



در گام نهایی هم این پیش بینی با چندین سیکل NEDC مختلف تست شد که خروجی به شرح زیر می‌باشد.

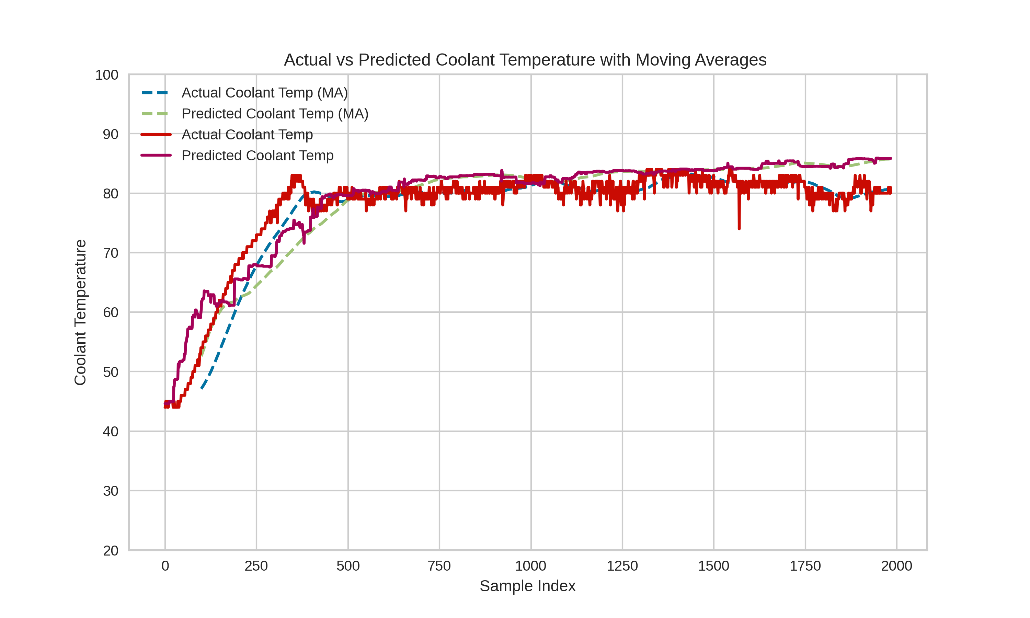


در ادامه برای تست و بررسی بهتر مدل 3 سفر که در آموزش دیده نشدخ بود را به صورتی انتخاب کردیم که دمای اولیه هر سفر متفاوت باشد یعنی دمای اولیه موتور متفاوت باشد. دمای اولیه موتور در سفر اول حدود 25 درجه است که در شکل زیر پیشبینی مدل را مشاهده می‌کنید.



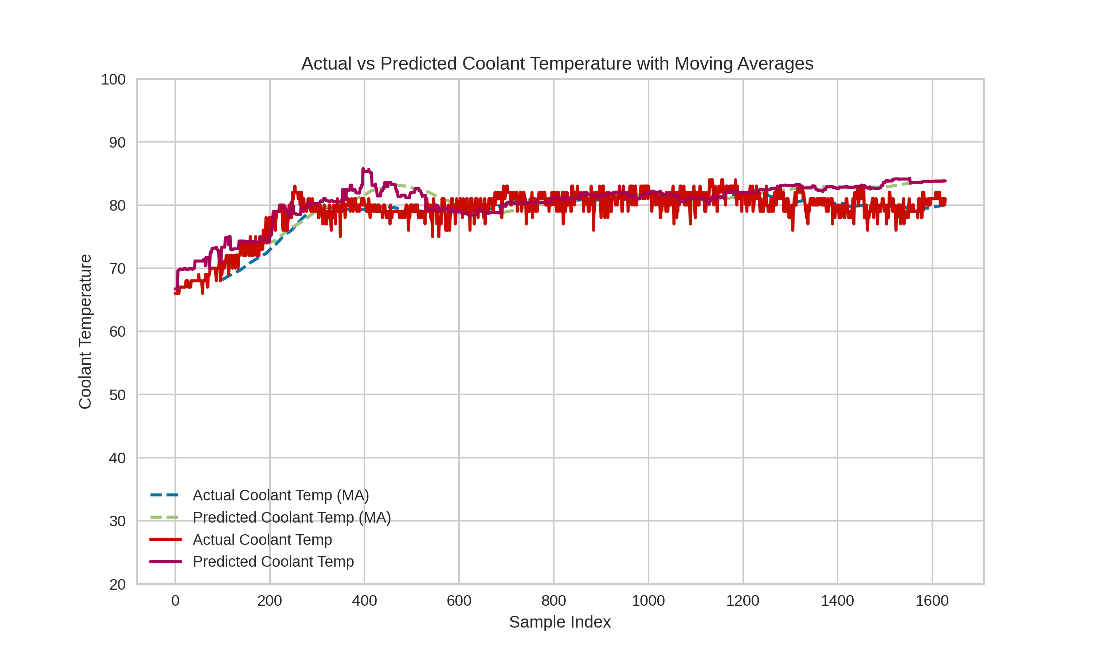
همانظور که در تصویر مشاهده می‌شود، زمان گرم شدن خودرو با دقت بالا پیشبینی شده همچنین در ادامه مسیر نیز با دقت خوبی دمای واقعی را پیشبینی کرده است.

برای تست بعدی سفری را انتخاب کردیم که دمای اولیه موتور حدود 45 درجه است. در شکل زیر نتایج این آزمایش قابل مشاهده است.

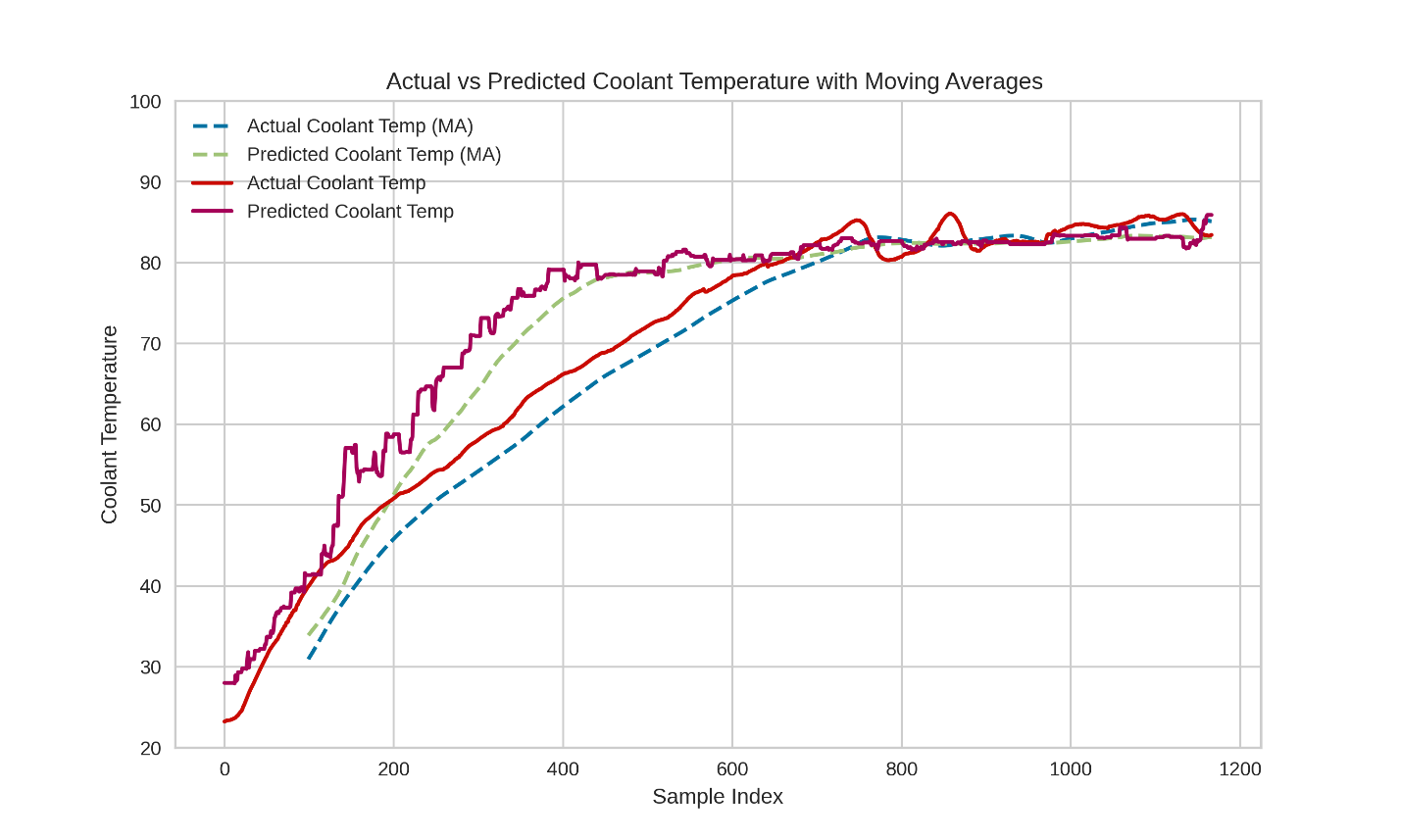


همانطور که در شکل بالا قابل مشاهده است، در این تست هم مدل دقت خوبی داشت و زمان و دما را به خوبی پیشبینی کرده است.

برای تست بعدی از دمای بالاتر استفاده کردیم در بتدای این سفر دمای موتور حدود 66 درجه است. نتایج پیشبینی در شکل زیر قابل مشاهده است.



در این تست هم مدل به خوبی عمل کرده و پیشبینی درستی از دمای موتور دارد. بعد از این تست ها، نوبت به تست مدل بر روی داده های NEDC رسیده است. نتایج این آزمایش در شکل زیر قابل مشاهده است.



همانظور که در تصویر بالا مشاهده می‌شود، دقت مدل به خوبی داده تست خودرو نیست، با بررسی‌های انجام شده توسط تیم علم داده عوامل ایجاد کننده خطا شامل : یکسان نبودن نوع جمع آوری داده و همچنین یکسان نبودن عوامل محیطی مانند دمای محیط است، چون داده‌های آموزشی این مدل در تابستان جمع آوری شده و در این فصل زمان گرم شدن موتور کمتر است همچنین روشن یا خاموش بودن سیستم تهویه خودرو مشخص نیست چون یکی از دلایل سریعتر گرم شدن موتور استفاده از تهویه است، از این رو برای نظر قطعی در این مورد باید داده‌های جدید در فصل‌های متفاوت دیده شود تا تاثیر عوامل محیطی بررسی شود.

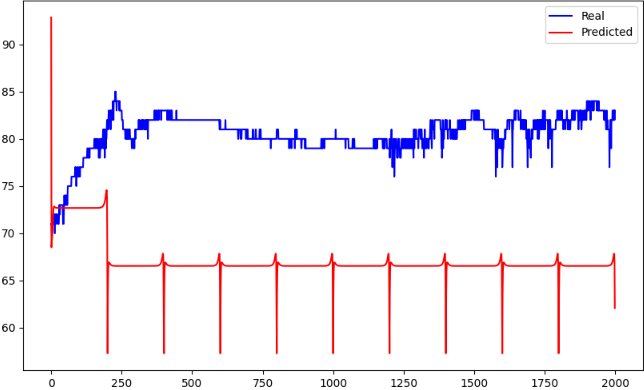
**نتیجه گیری**

هدف ما در این بخش بررسی کارای الگوریتم‌های XGBoost و CNN و BiLSTM برای پیشبینی دمای خودرو است.

ابتدا یک فایل برای تست انتخاب کرده و از داده‌های یکسان برای آموزش استفاده کردیم، ویژگی‌های مورد استفاده در این بخش time, Vehicle\_Speed, Current\_gear, initial\_coolant\_temp بودند و مقادیر نرمال سازی برای هر ویژگی به ترتیب time = 0 to 1080798 و0 to 137 Vehicle\_Speed = و Gear = 0 to 6 و initial\_coolant = 18 to 102 است.

قبل از پیش پردازش به هر یک از این ویژگی‌ها به اندازه 10درصد از انحراف معیار آن‌ها برای داده افزایی نویز اضافه شده است.

بعد از این پیش پردازش‌های لازم ابتدا از شبکه BiLSTM که مورد استفاده در مصرف سوخت بود استفاده گردید و نتایج به دلیل کمبود داده خوب نبود و شبکه آموزش ندید که در تصویر زیر قابل مشاهده است:



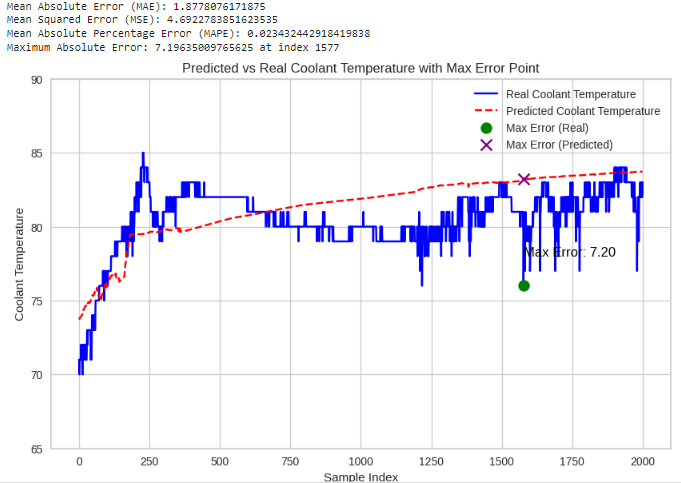
که همانظور که مشاهده میشود این روش برای انجام این کار مناسب نیست.

بعد از این به سراغ شبکه CNN رفتیم که از دو ساختار local و global ساخته شده است این دو ساختار به این معنی است که شبکه اول که local است که ویژگی‌های پنجره به اندازه 5 را استخراج میکند و global ویژگی‌های پنجره به اندازه 20 را استخراج میکند و دید کلی‌تری دادرد و در اخر این دو با یکدیگر ترکیب شده و به یک FCN یا شبکه ساده داده میشوند و خظای مورد استفاده برای آموزش مجموع مربعات خطا یا MSE است.

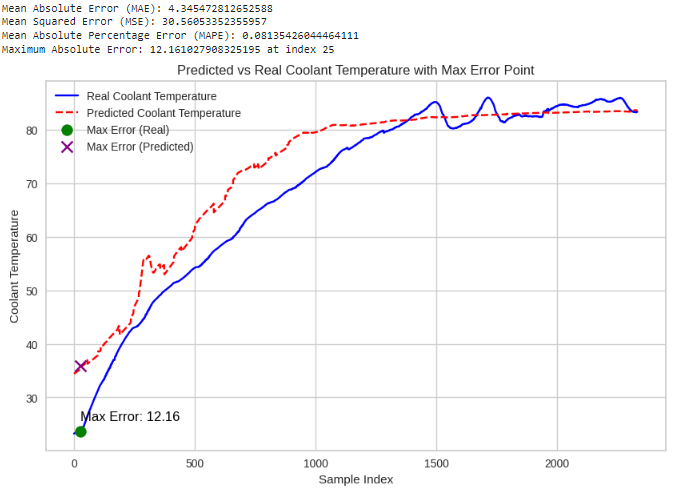
برای تست این شبکه از یک سفر ثابت و سیکل NEDC خودروی Dena\_TC+\_MT استفاده شده است. این دو تست برای شبکه XGBoost نیز ثابت است و تغییری ندارد.

تست انجام شده بر روی سفر دیده نشده برای شبکه CNN به صورت زیر است که اگر MAE را معیار قراردهیم یک میانگین خطای 1.8درجه سانتیگراد دارد یعنی اگر دمای واقعی 80 درجه باشد پیشبینی این شبکه یا 81.8 درجه و یا 78.2 است و خطای میانگین یا MAPE برابر 2درصد است.

بیشترین خطا نیز 7 درجه است که در تصویر زیر قابل مشاهده است.

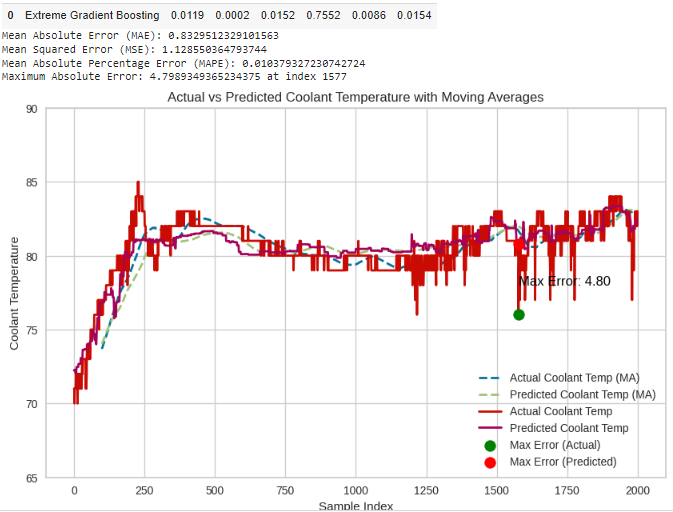


بعد از این نوبت به تست NEDC میرسد برای این تست نتایج به صورت زیر است که میانگین خطا برابر 4 است یعنی به صورت میانگین اگر دما 80 درجه باشد پیشبینی ما 84 یا 76 است. همچنین درصد خطا برابر 8درصد است و همچنین بیشترین خطا برابر 12 در نقطه نشان داده شده است.

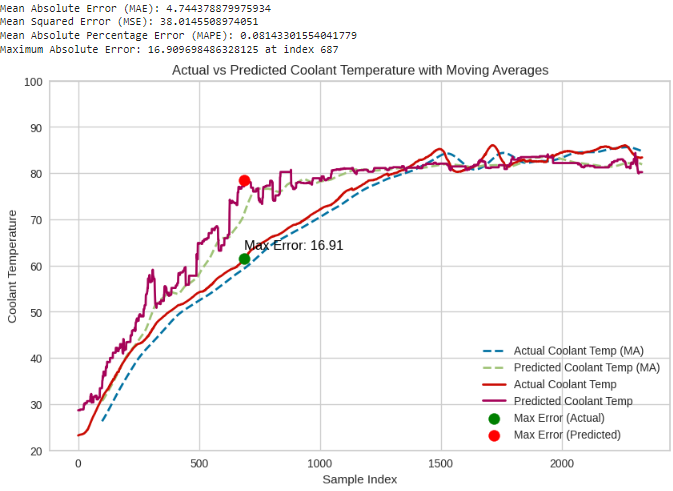


در این مرحله ما از XGBoost برای آموزش و پیشبینی استفاده کردیم. در این مرحله هم مانند مراحل قبل داده‌های آموزشی یکسان است و همچنین برای تست از یک سفر و NEDC ثابت استفاده شده است.

در این قسمت ابتدا تست بر روی سفر انجام شده همانطور که در شکل زیر مشاهده میکنید میانگین خطا MAE برابر 0.8 است یعنی اگر دمای واقعی برابر 80 درجه باشد پیشبینی مدل برابر 80.8 یا 79.2 خواهد بود و همچنین درصد خطا برابر 1درصد است که دقت خوبی دارد و زمان آموزش بسیار کمتری نسبت به CNN لازم دارد و همچنین پیچیدگی تنظیم پارامتر CNN را در این بخش نداریم. بیشترین خطا نیز 4 درجه و در شکل نمایش داده شده است.

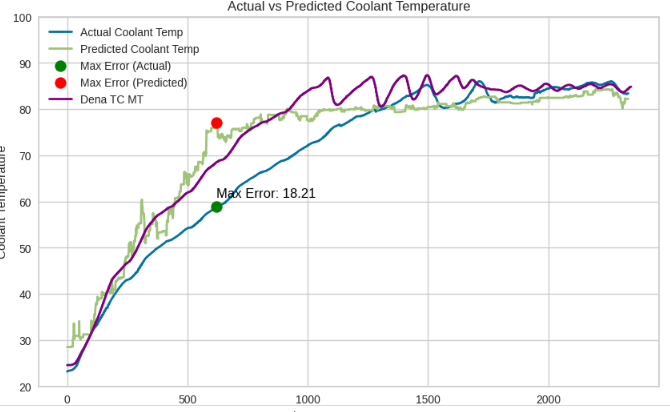


بعد از تست سفر نوبت به تست و آزمایش بر روی NEDC است که میانگین خطا در این بخش برابر 4.7 است یعنی اگر دمای واقعی 80 درجه باشد مدل ما دما را 84.7 یا 75.3 پیشبینی میکند و میانگین خطا برابر 8 درصد است همچنین بیشترین خطای ایجاد شده برابر 16.9 درجه در نقطه مورد نظر نمایش داده شده است.

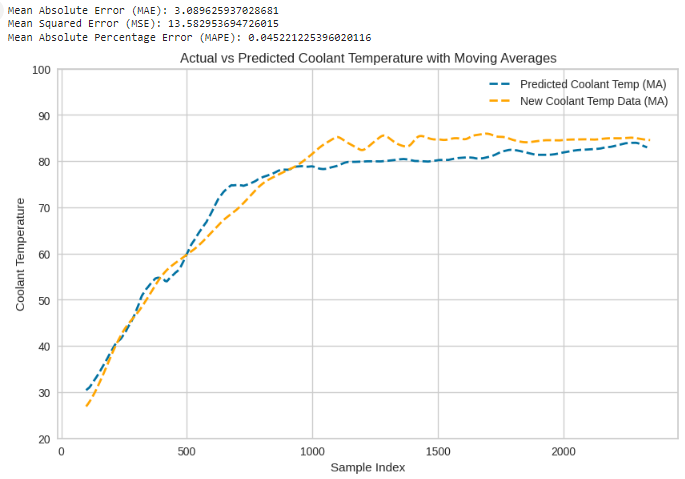


برای نتیجه گیری کلی بر اساس نتایج به نظر می‌رسد با مدل ساده‌تر XGBoost ما به هدف اوله خود یعنی پیشبینی دما و استفاده آن به عنوان ورودی برای آموزش و پیشبینی مصرف سوخت با دقت خوبی انجام شده‌است البته برای NEDC هم دلایل خطا در بالا گفته سده که عبارت است از دمای محیط داده‌های آموزش و جنس داده برداری متفاوت، اما در مجموع استفاده از XGBoost درست است و با داده‌های کمتر هم مشکلی ایجاد نمی‌شود.

بعد از این بررسی ما NEDC خودروی 955 که TC جاری هست را وارد نمودار کرده ولی چون اطلاعات دنده آن موجود نبود از پیشبینی TC+ استفاده کردیم ولی تظبیق نمودار انجام شد که نتیجه به شکل زیر است.



که خط بنفش دمای NEDC خودروی DENA TC یعنی خودروی 955 استو خط سبز پیشبینی ما با اطلاعات TC+ است ولی همانطور که مشاهده می‌شود دقت خوبی دارد همچنین ما خطارا برای این مدل هم حساب کردیم و میانگین مجموع خطا MAE برابر 3 شد یعنی اگر دمای واقعی برابر 80 باشد پیشبینی ما 83 یا 77 است. همچنین MAPE درصد خطا برابر 4.5درصد است.



اموزش قبل از این تست برابر با 13هزار کیلومتر بود که برای ارزیابی کارایی XGBoost ان را به 656 کیلیومتر ولی با اطلاعات جامع یعنی این اطلاعات نماینده مناسبی برای کل سفرها بود را کاهش دادیم و نتایج این NEDC به صورت تصویر بالا شد و حتی کمی دقت بهتری پیدا کردیم. یعنی با داده های کم هم میتوان این پیشبینی را انجام داد.