MACHINE LEARNING EXAM 1

Tursunov Akhmadjon

**1 Model Reporting**

* Ushbu bo'lim loyihaning qisqacha tavsifi, maqsadi va asosiy natijalarini taqdim etadi.
* Maqsad:
* Ushbu loyiha maqsadi taxi mijozlaridan olingan ma’lumotlarga asoslanib, asoslanib yo’l haqini aniqlashdir.

**Features Haqida Umumiy Ma'lumot**

Ma'lumotlar to'plami 1 000 tani tashkil etib, 11 ta xususiyat(features)dan iborat. Target feature “Trip\_price”.

* Raqamli o'zgaruvchilar: Trip\_Distance\_km, Passenger\_Count, Base\_Fare, Per\_Km\_Rate, Per\_Minute\_Rate, Trip\_Duration\_Minutes, Trip\_Price.
* Kategorik o'zgaruvchilar**:** Time\_of\_Day, Day\_of\_Week, Traffic\_Conditions, Weather.

**Data Preprocessing jarayoni**

* Missing value bilan ishlash:

Target featureimizdagai Null valuelarni mode() orqali to’ldirdik

* Coding:

Random forest uchun kategorik o'zgaruvchilarni barchasini label encoding bilan raqamli ko’rinishga o’tkazdik. Bu data setimizni toliqligicha numerical o’rinishga o’tkazishga yordam beradi.

* Scaling:

Malumotlarimizni Standard scaling orqali muvozanatga olib keldik ammo bu faqat Random forest uchun kerak.

* Maʼlumotlarni boʻlish:

Dataset imizni 80% ini traingn uchun qolgan 20% izini test uchun bo’lib oldik

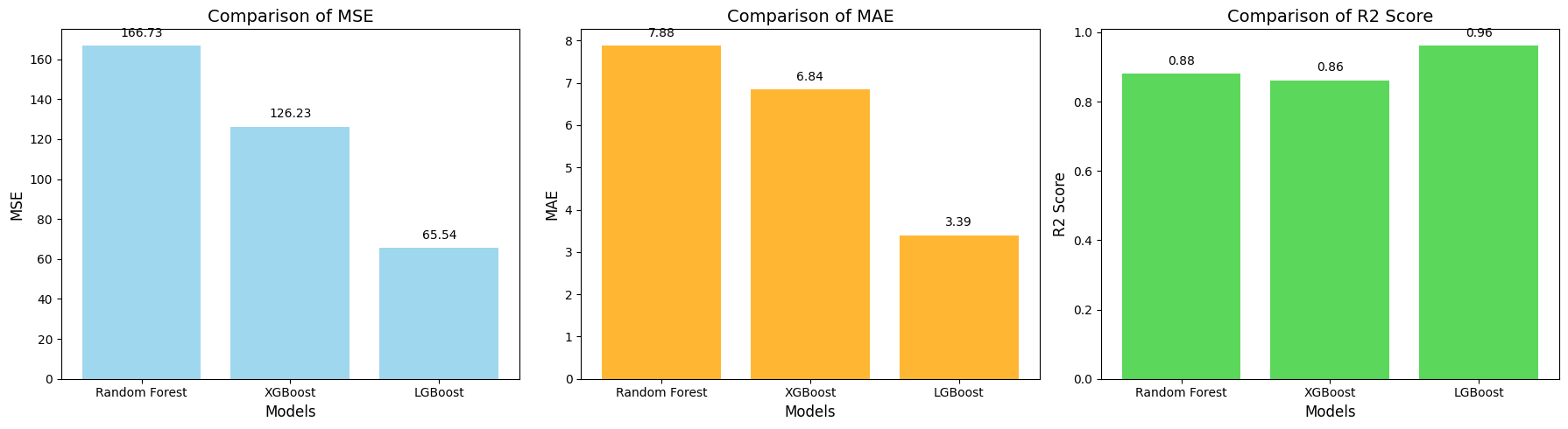
**Modelni Rivojlantirish**

* Tayyor data set imizni XGBoost, LGBoost va Random Forest modellar ga o’qitib, ularni o’zaro solishtiramiz.
* XGBoost, LGBoost da data preprocessiong jarayoni ancha kam bo’ldi va ularning qulay tomoni hisoblanadi.

Modelni Baholash

* Mean square error
* Mean absolute error
* R2 score

Xulosa

****

Natijamiz quyidagicha eng yaxshi model LGBoosting bo’ldi u har bir baholash parametrlarida qolganlariga nisbatan yaxshi ko’rsatgich ko’rsatdi.