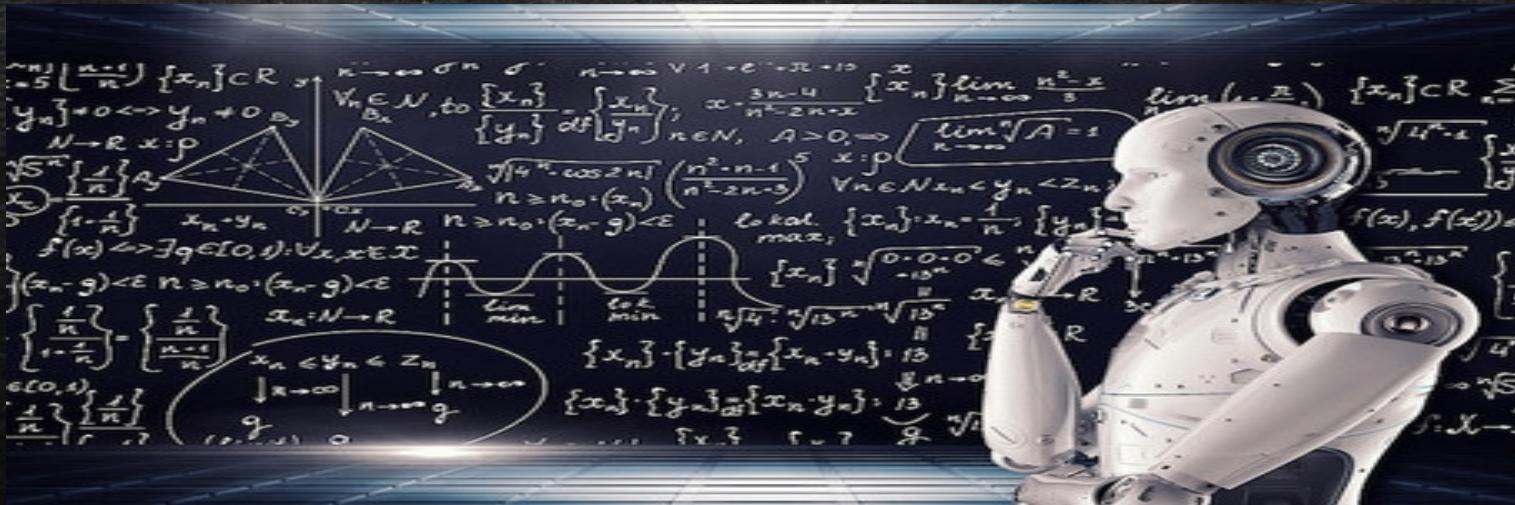


Machine Learning and Content Analytics Recommendation System



Professors: H. Papageorgiou, G. Perakis

Students: G.Vlassi p2822001

S.Vretteas p2822003

A.Ntetsika p2822014

D.Mentakis p2822024

Semester Project

Why

... use Machine Learning?



What

... is Machine Learning?



Contents

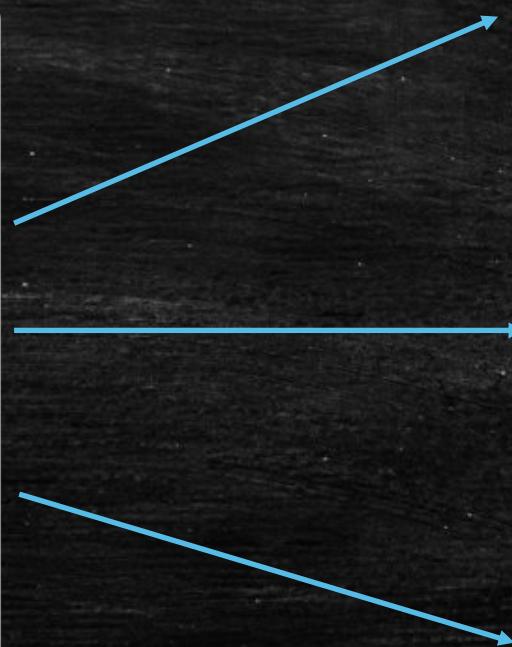
- Scope
- Approach Overview
- Data Analysis
- Modeling
- Hyperparameter Tuning
- Results
- Problems, Improvements and Future Use



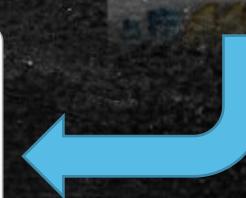
Scope

Greek pay TV based on NETFLIX

- Recommendation
- Predicting

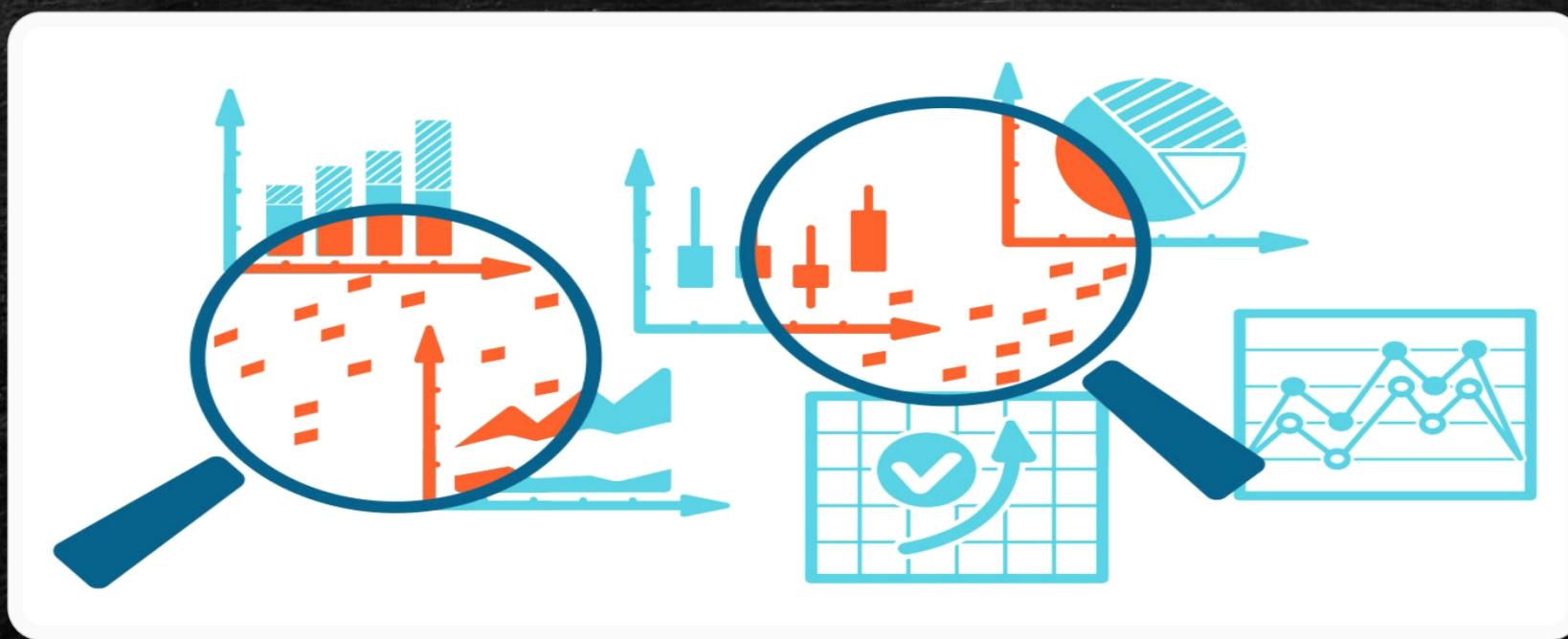


Approach Overview



Data Extraction Process

- Descriptive Analysis
- Explanatory Data Analysis



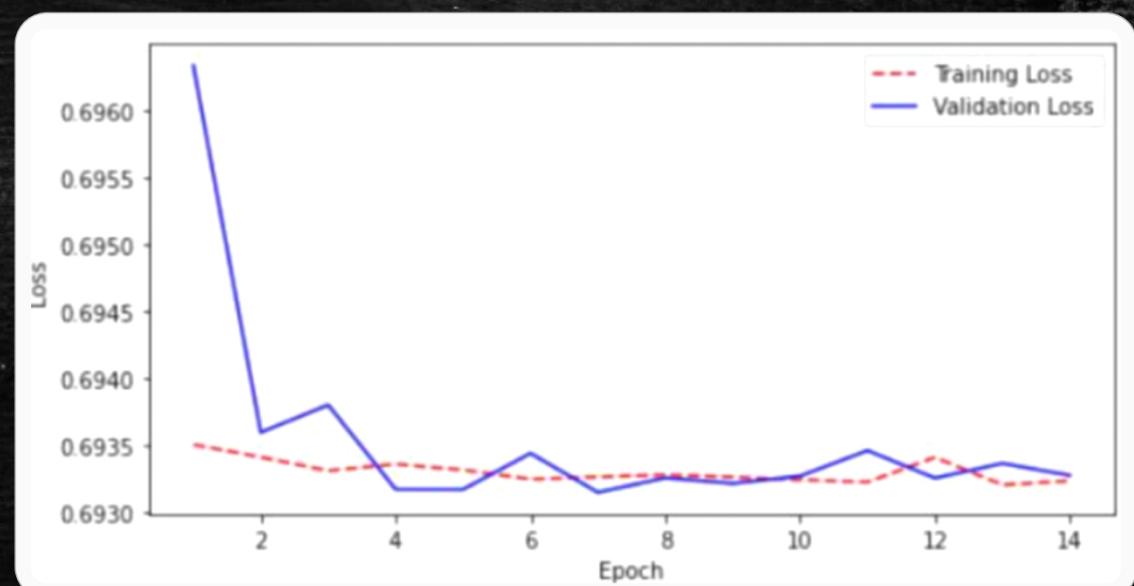
Model 1: Predict User's Gender from Username



Model 1: Predict User's Gender from Username

- LSTM

Hyperparameter	Values
Epoch and batch size	50 and 64
Optimizer	Adam
Function used	Sigmoid
Learning rate	0.002



Model 2: What's Next Prediction



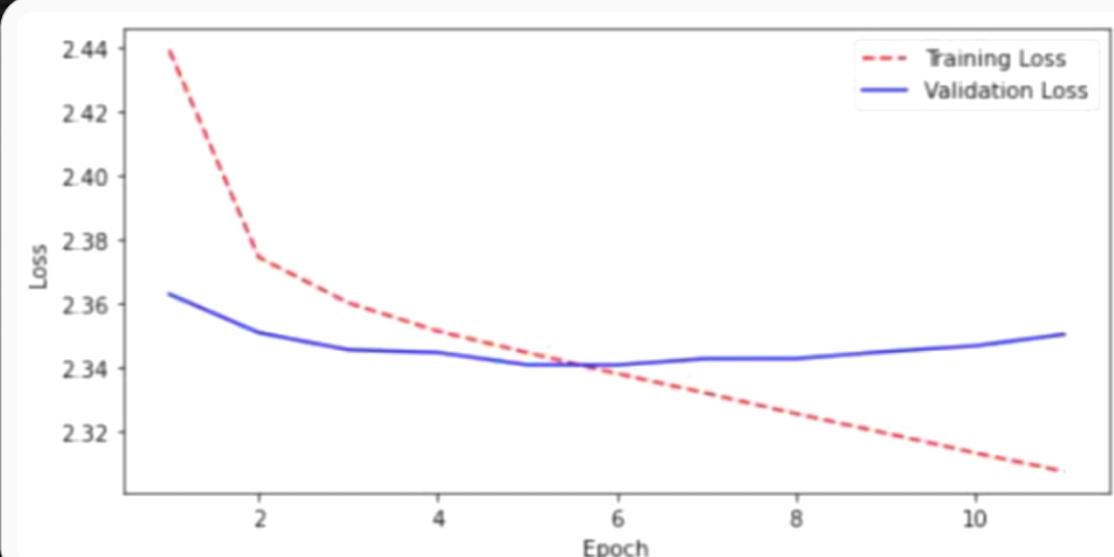
Don't forget
your vocabulary
😊



Model 2: What's Next Prediction

- LSTM

Hyperparameter	Values
Epoch and batch size	1002 and 256
Optimizer	Adam
Function used	SoftMax
Learning rate	0.002



Model 3: Collaborative Filtering

I want only
those kind of
movies



Model 3: Collaborative Filtering



ASSET_NAME	Correlation	total ratings
THE COMMUNE	0.9997631224347492	119
TELLE MERE, TELLE FILLE	0.9997631224347492	426
BACK TO BURGUNDY	0.9997631224347492	210
WHEN THE CITY DIES	0.9997631224347492	114
VALERIAN	0.9997631224347492	489
ACTS OF VENGEANCE	0.9997631224347492	10896
TWO IS A FAMILY	0.9997631224347492	359
SEE YOU UP THERE	0.9997631224347492	229
LES HOMMES DU FEU	0.9997631224347492	129
RETURN TO MONTAUK	0.9997631224347492	149

Model 4: Multi-Model

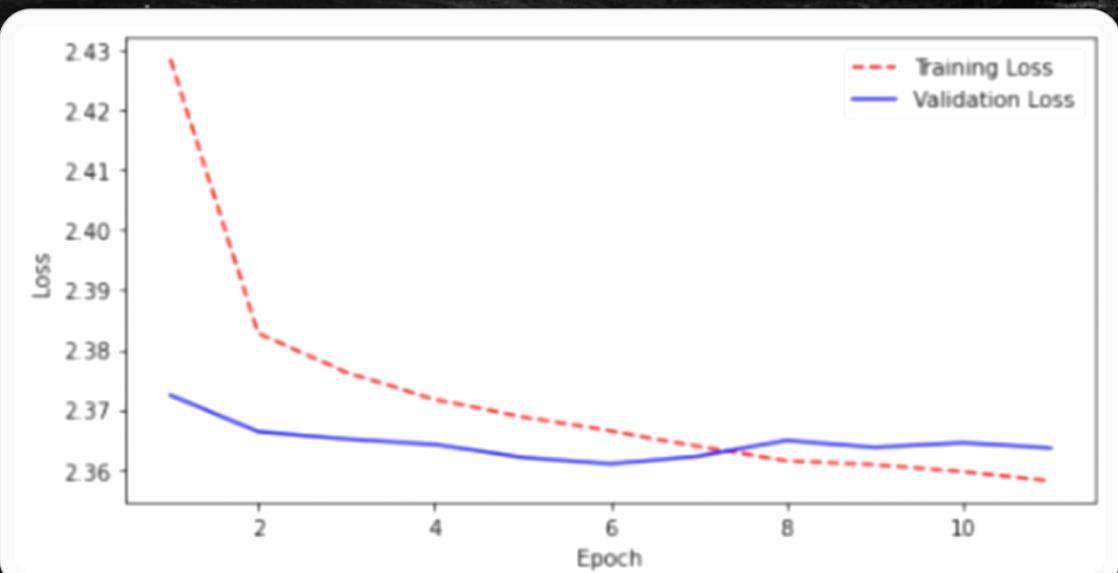
- And if we know the gender of the user ?
- Can we find the next genre that user will watch?!



Model 4: Multi-Model

- LSTM

Hyperparameter	Values
Epoch and batch size	100 and 256
Optimizer	Adam
Function used	SoftMax
Learning rate	0.002



Model Comparison

	Model 1	Model 2	Model 3	Model 4
Test loss	0.6931	2.3455	X	2.3638
Accuracy	0.5021	0.4083	X	0.4078
Stop after epochs	21	11	X	11
Correlation	X	X	Each row	X

Problems

- Dataset for specific year
- Machine problem (need better pc with more RAM , lost time by finding google collab)
- Github (compile packages)

Future Use as an Application

- Companies

New company or start up try his luck against nova, ote, ertflix

- Sports Club

From viewers predict their opinions and create package friendly to them to watch their favorite teams



QUESTIONS?