Bloom's Taxonomy ML Classifier

Neha Vijay Patil

Defining the problem

Problem Definition

BLOOM'S TAXONOMY – COGNITIVE DOMAIN (2001)

HIGHER-ORDER THINKING SKILLS

CREATING
Use information to create something new

EVALUATING
Examine information and make judgments

ANALYZING
Take apart the known and identify relationships

APPLYING
Use information in a new (but similar) situation

UNDERSTANDING
Grasp meaning of instructional materials

REMEMBERING
Recall specific facts

LOWER-ORDER THINKING SKILLS

Facilitating the creation of balanced question papers to evaluate students on different levels based on Bloom's Taxonomy.

Problem Definition

- Existing method of implementation: Simple mapping of verbs -> levels

Limitations:

- action verbs are not always the indicator of level. (what is...)
- some verbs are associated with multiple levels. (describe)

Proposed solution:

- ML Classifier will solve the above issues and automates the categorisation process

Literature Survey

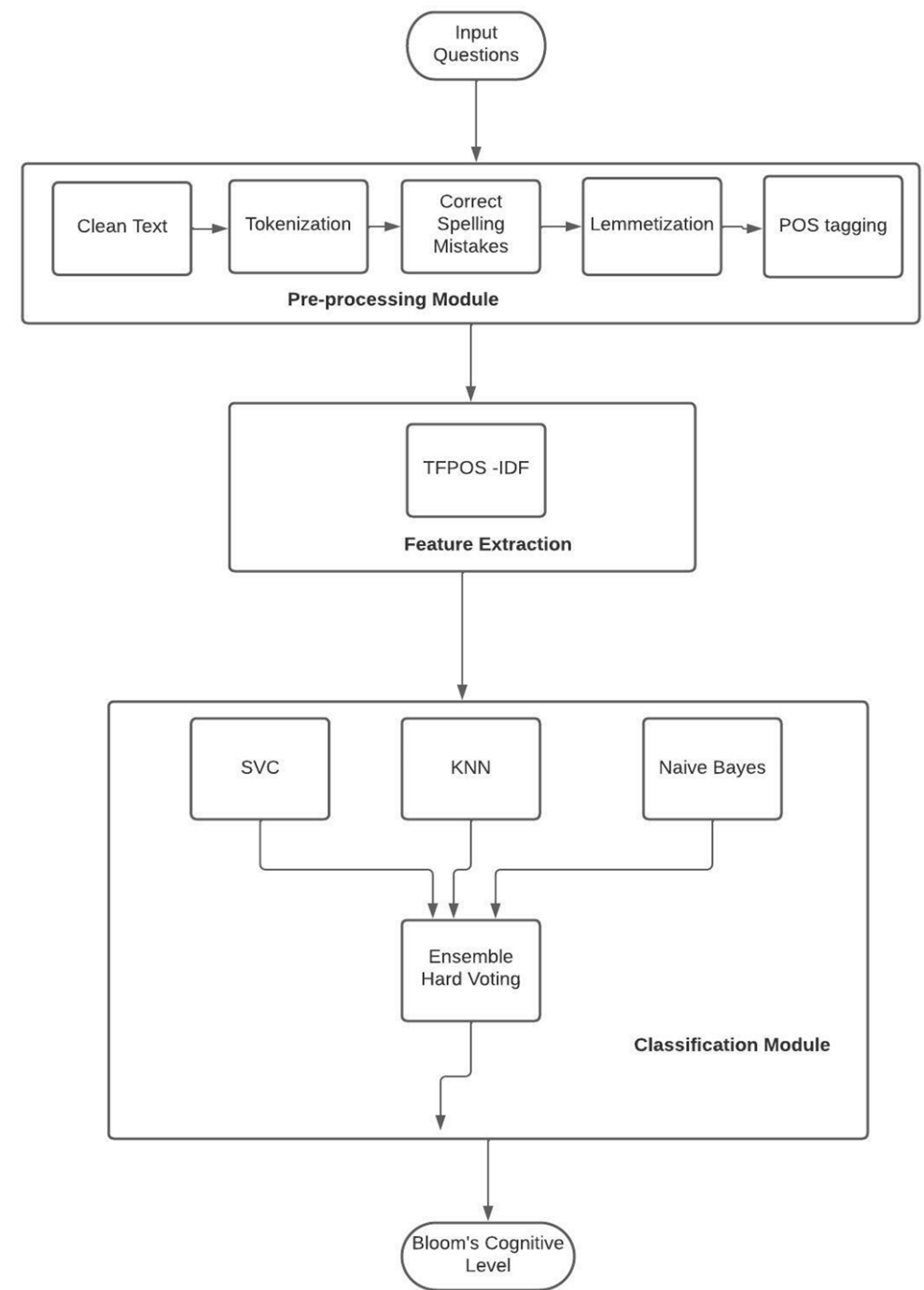| Papers	[11]	[12]	[13]	[9]	[15]	[16]	Our Algorithm
Feature extraction	NA	Chi-Square+Laplace Smoothing	Chi-Square, Mutual Information, and Odd Ratio	TFPOS-IDF and word2vec	NA	Pos Tagging and Word Vector	TFPOS-IDF
Classifiers	SVM, K-NN.	Naive Bayes	Naïve Bayes(NB), (KNN).	KNN, Logistic Regression, SVM	SVM-Light	Rule based grammar, , Ensemble Technique on KNN, SVC and NB, Wordnet.	Ensemble Technique on KNN, SVC and NB.
Accuracy	SVM: 69% K-NN: 65%	60.63%.	KNN: 87%. NB: 83%.	KNN: 71.1%, LR: 82.3% SVM: 83.7%	87.4%. Poor recall and F-measure values	Ensemble: 82.5% Rule Based: 60%	93.5% Precision: 94% Recall: 93%

Data & Model Design

Snippet of the dataset

4]:

	Text	Target
1221	name five cities in us	remembering
588	revise the story and specify the type of audie...	creating
302	analyze the movements and sounds of a frog	analyzing
37	list reserved words in c programming	remembering
281	what approach would you use to prove the truth...	applying
403	can you develop a proposal which would	evaluating
547	given the data weve looked at on this topic ev...	creating
530	determine if a character's actions were heroic	creating
332	examine what helps to make a good olympics thi...	analyzing
25	identify the correct definition of osmosis	remembering

Distribution of levels in dataset

Blooms Taxonomy Levels Distribution
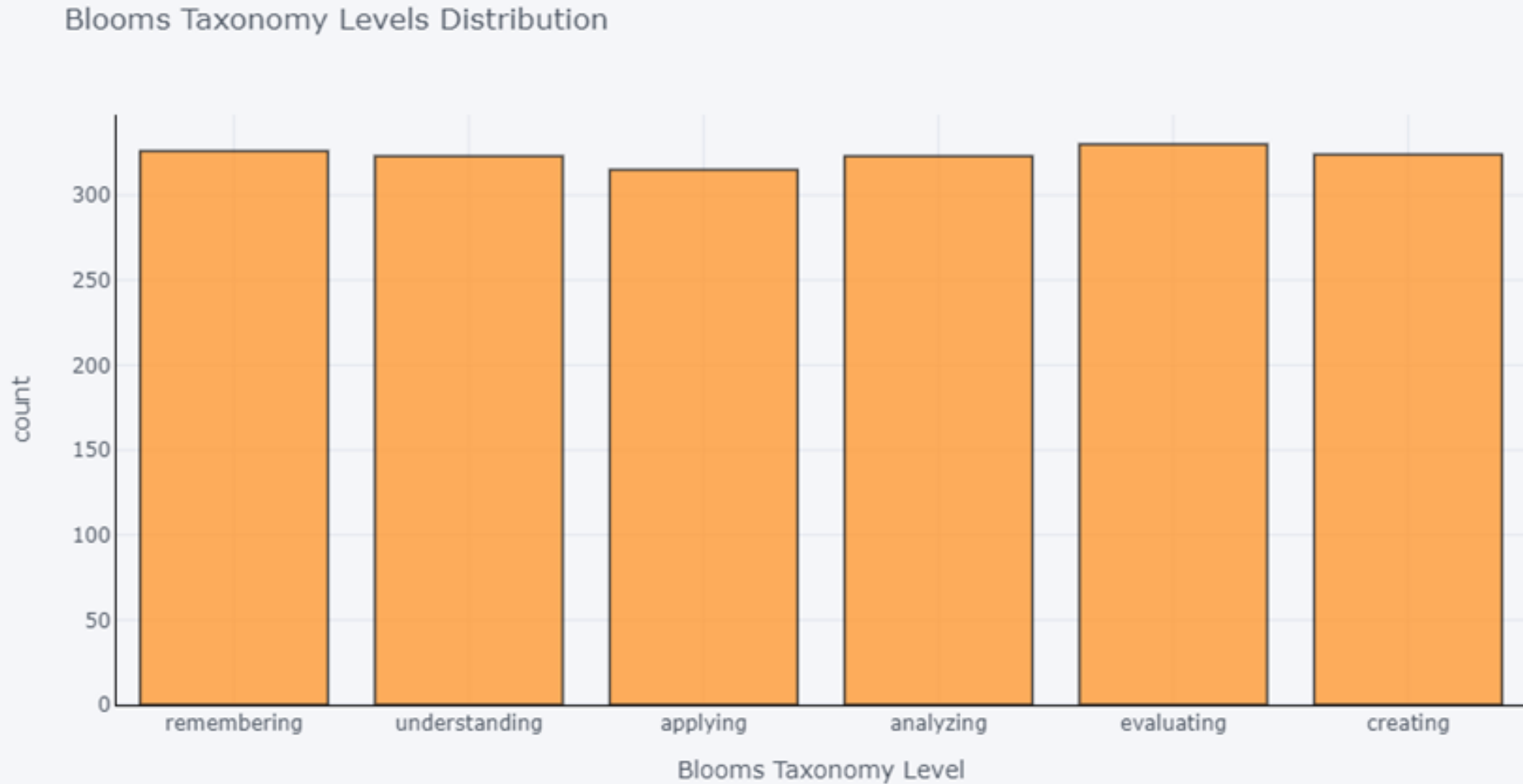
Model Design

- Pre-processing

- Customised feature extraction algorithm to give high weightage based on positioning.

- Ensemble ML classifier combining results of 3 models: KNN, SVC and Naïve Bayes using hard voting.
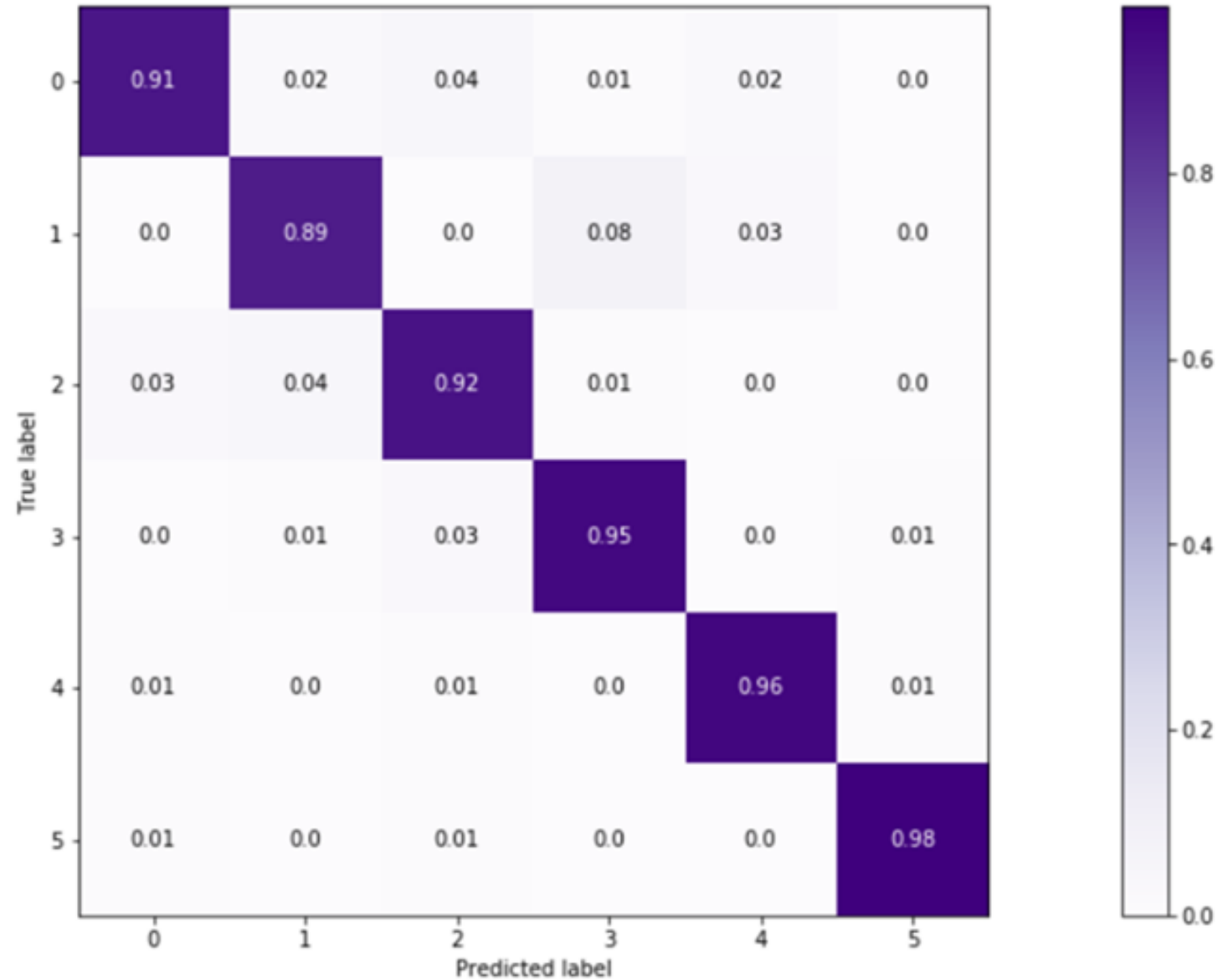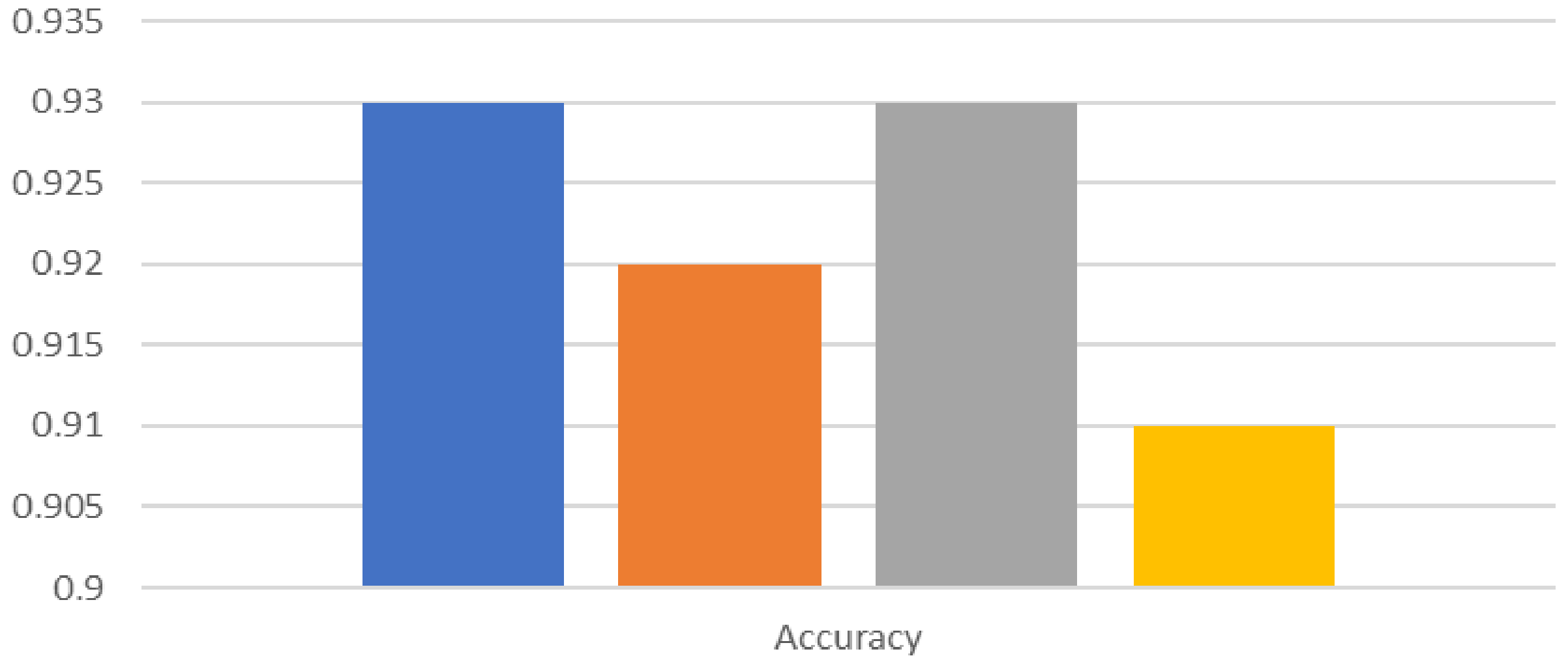
Implementation

https://github.com/neha7799/BloomsTaxonomyClassifier/blob/main/Project%20-%2027march-%20model%20presentation.ipynb
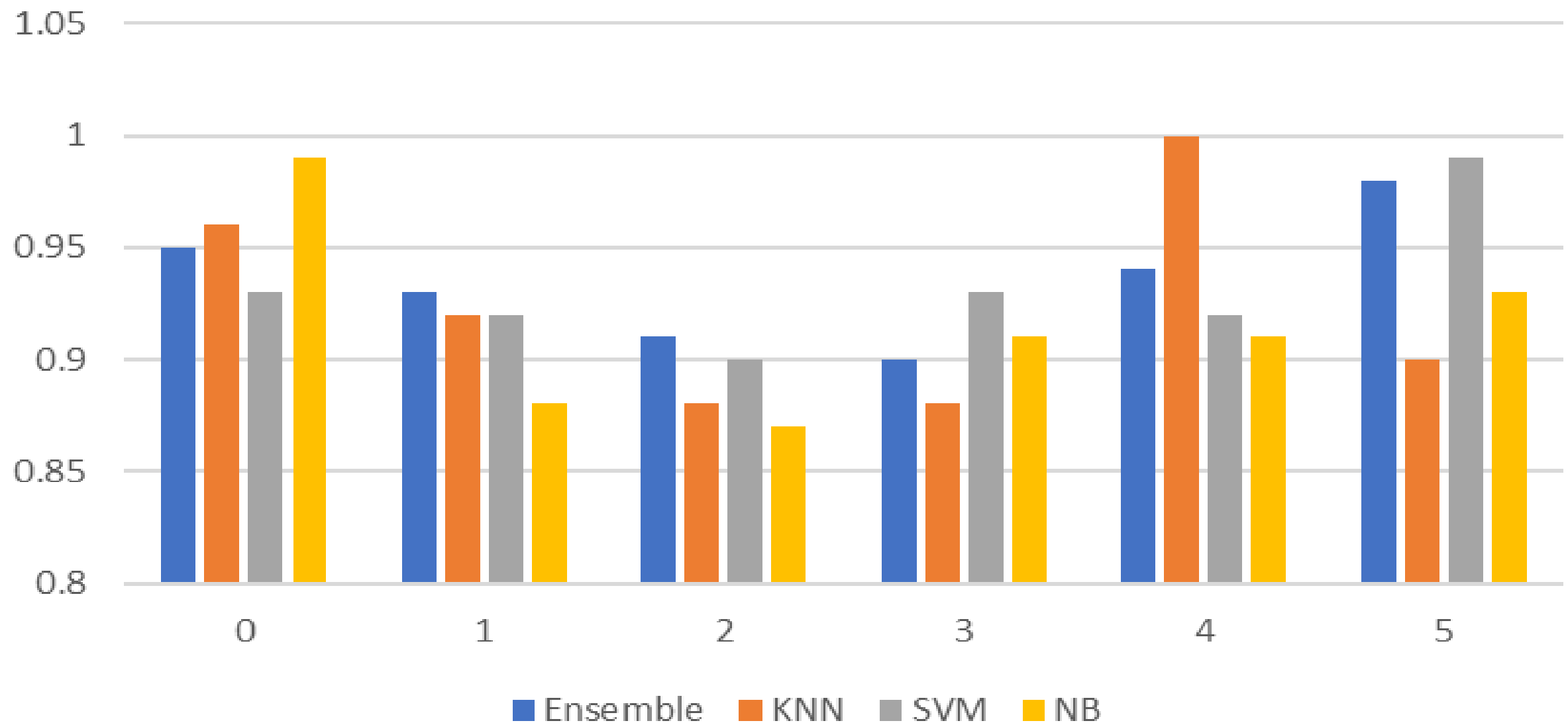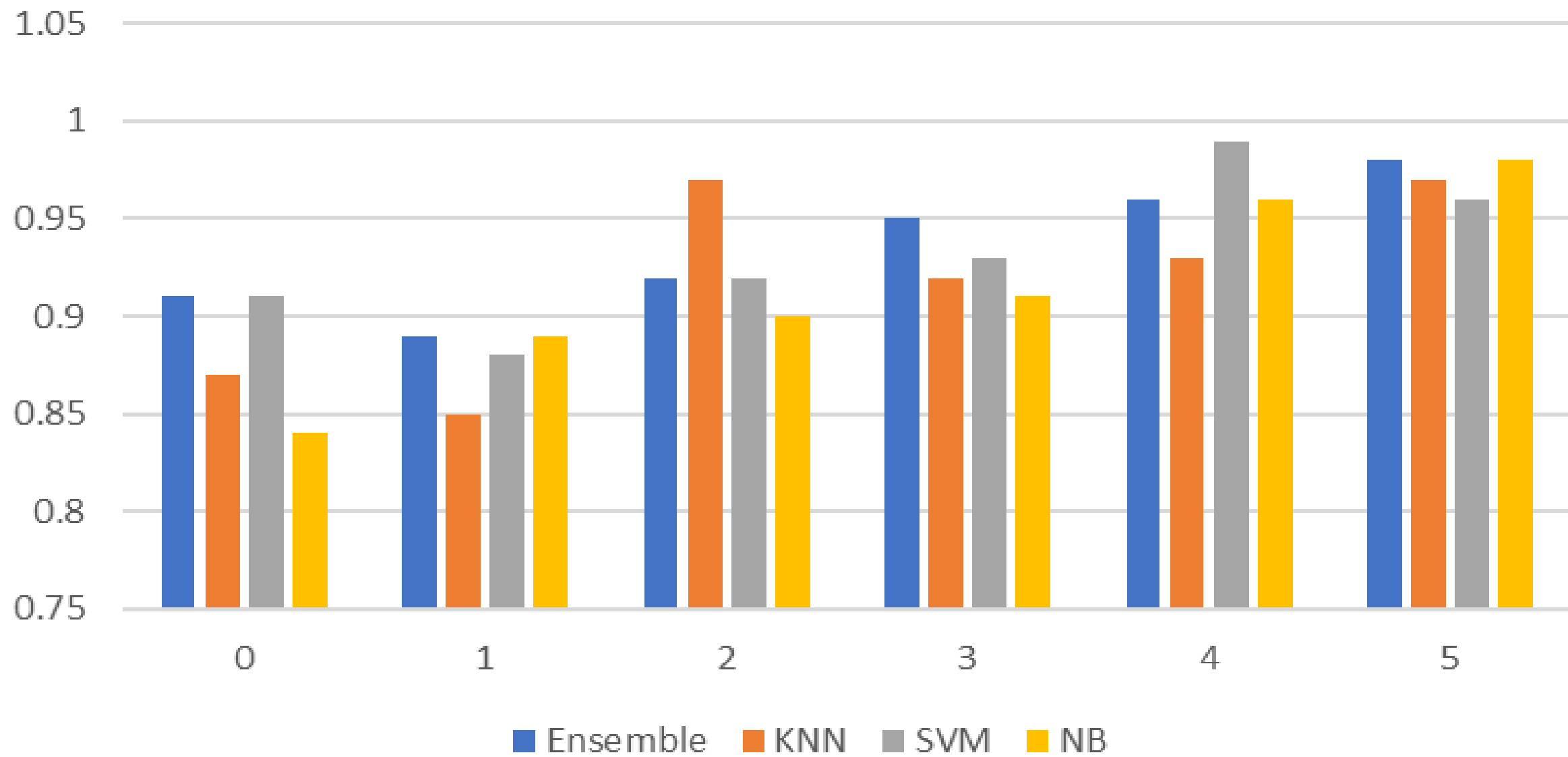
Results

Confusion Matrix

Accuracy

Precision

Recall

F1 score

Ensemble   KNN   SVM   NB
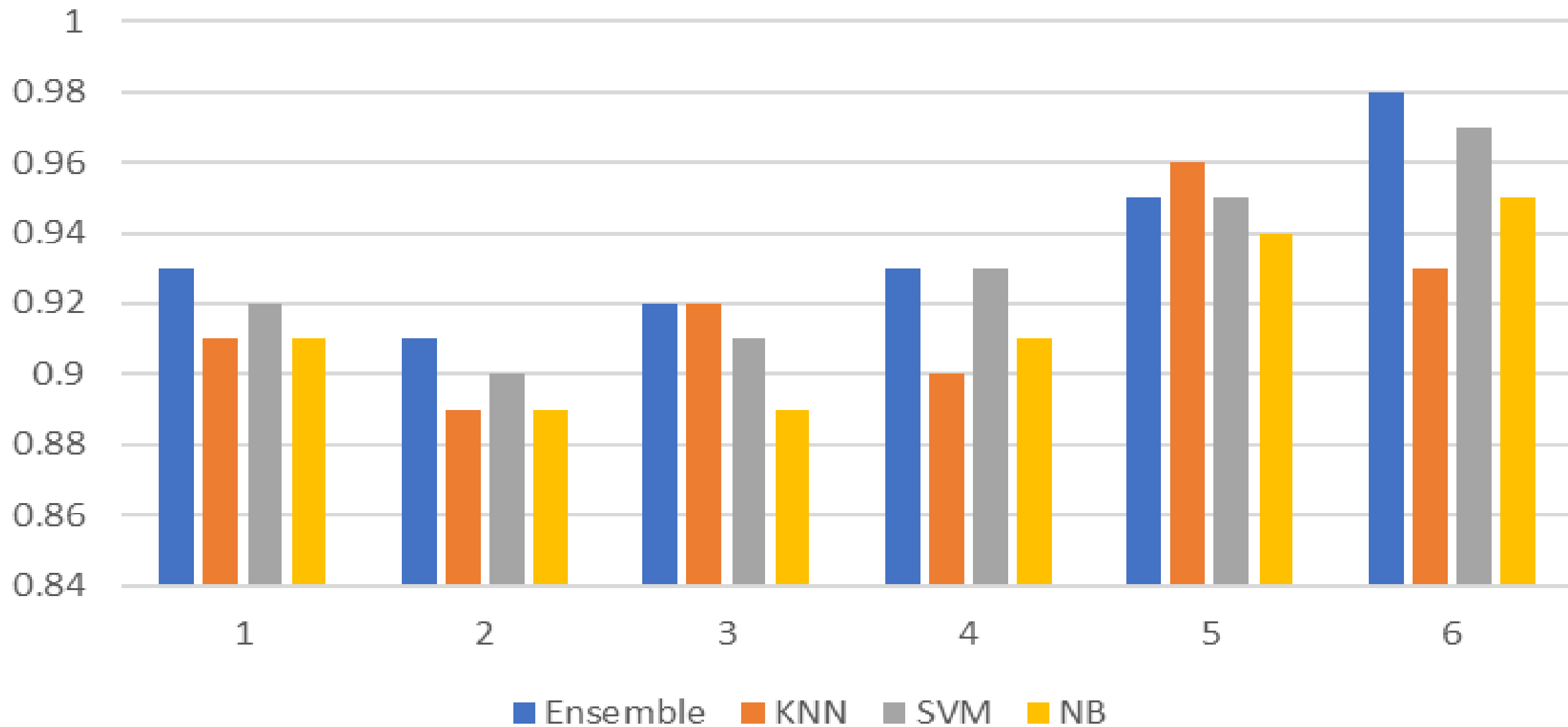
Future Work

- More data collection.
- Apply cross validation techniques.
- Check for overfitting.

More about me

- Cloud Infrastructure Intern (IAM Team) : Azure, PowerShell scripting
- Research Assistant : Software Engineering & Human Computer Interaction
- Full Stack Development course and projects: Angular
- Data Analytics Intern : Python, REST API, MySQL, XML
- CyberSecurity Research Intern: Reverse Engineering, Linux, Vulnerability, Malware analysis

Thank you!
Open to questions