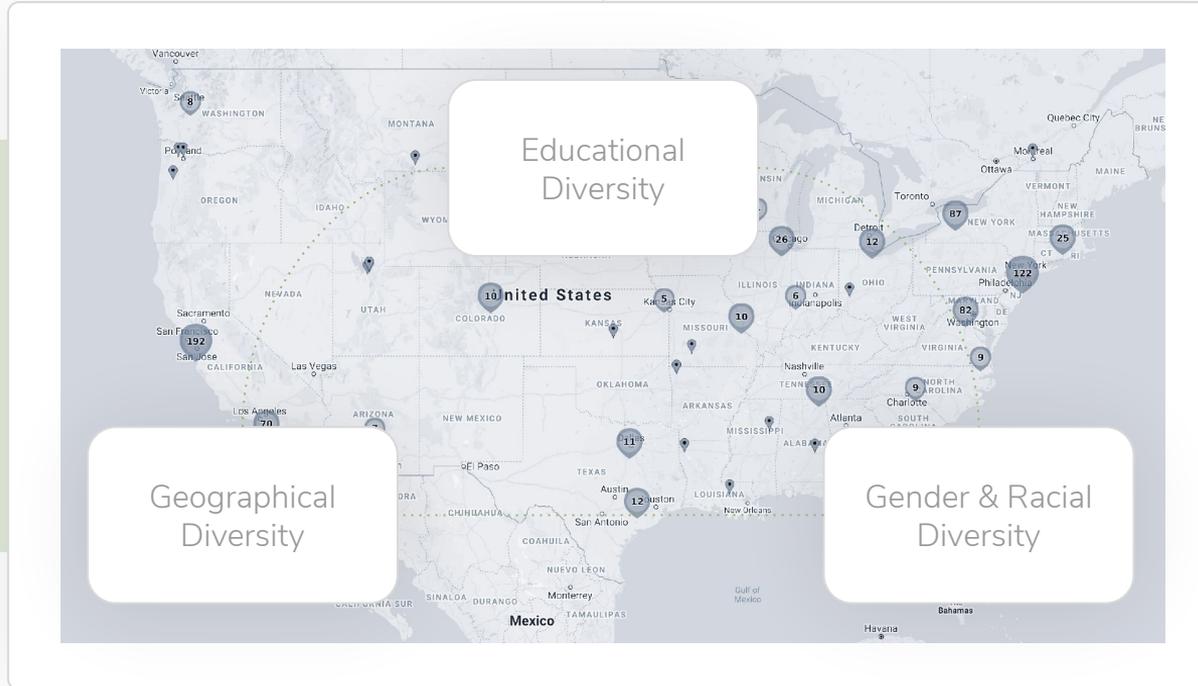


SA
STEM-AWAY

STEM-Away®
Machine Learning Pathway



Rethinking internships as the way to expand the work-ready STEM talent pool beyond traditional demographics and colleges



STEM-Away Virtual-Internships are organized as Mentor Chains® Projects

STEM-Away Mentors
(STEM Professionals)

Mentor Chains® Leads

Mentor Chains®
Participants

Mentor Chains® Observers

STEM-Away® Virtual-Internships are organized as Mentor Chains® Projects:

Industry experts mentor student leads.

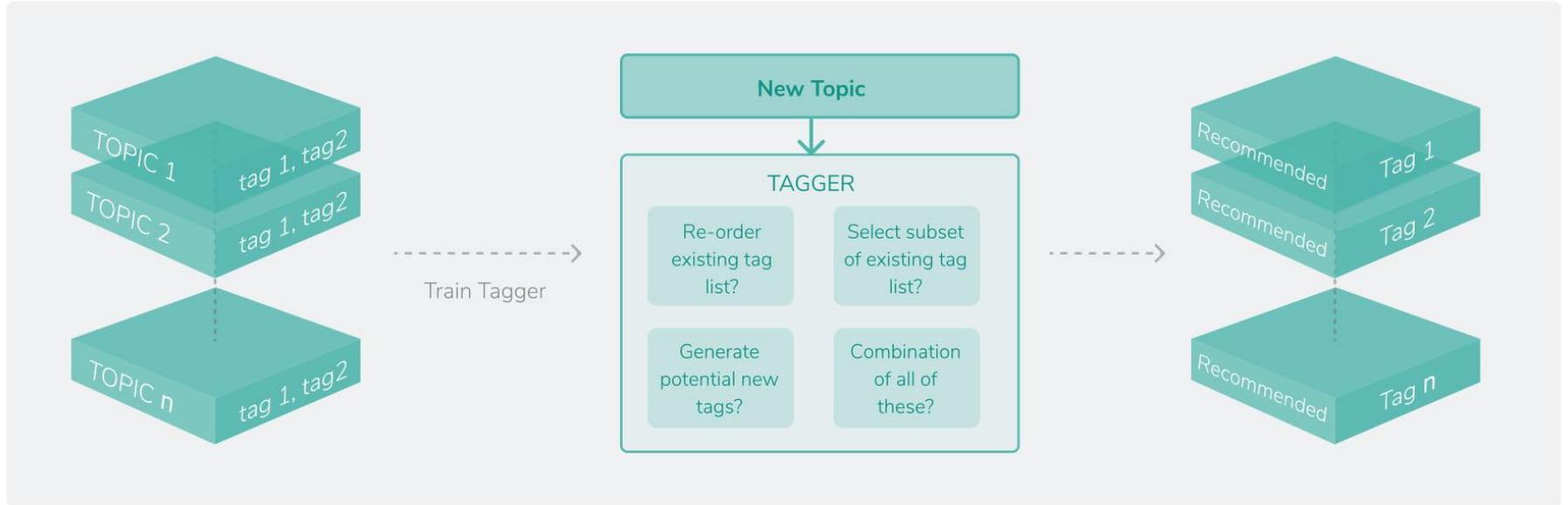
Leads mentor teams of 8-15 students.

Granularity of Mentor Chains® dependent on pathway.

Exponential increase in STEM internship opportunities with the Mentor Chains® structure.

Building the next generation of inclusive STEM leaders by giving students an early start on leadership & mentoring skills.

A Level 2 Project: NLP Annotation Tool with Active Learning

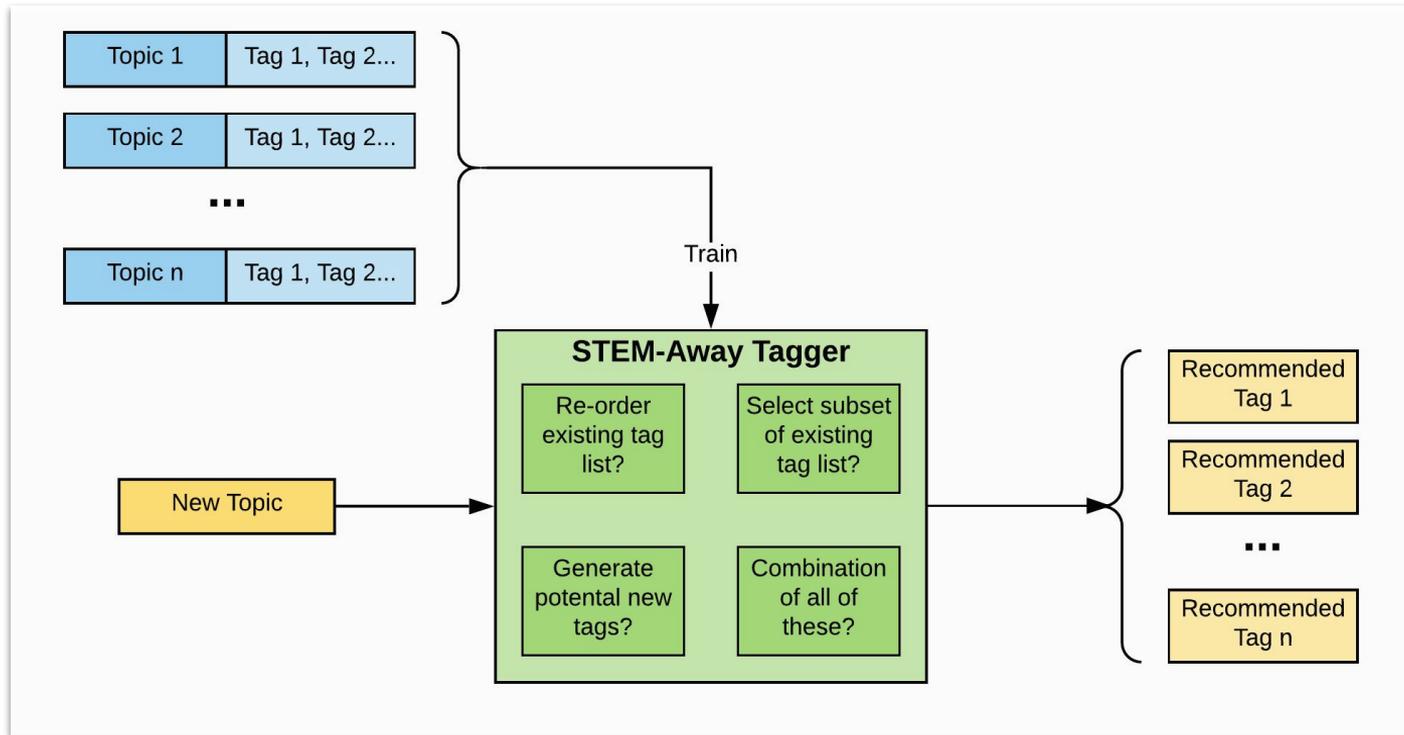


Agenda

1. Project Outline
 2. Software Architecture
 3. Hyperparameters, Metrics,
& Benchmarking Results
 4. Soft Skills
 5. Questions and Discussion
-

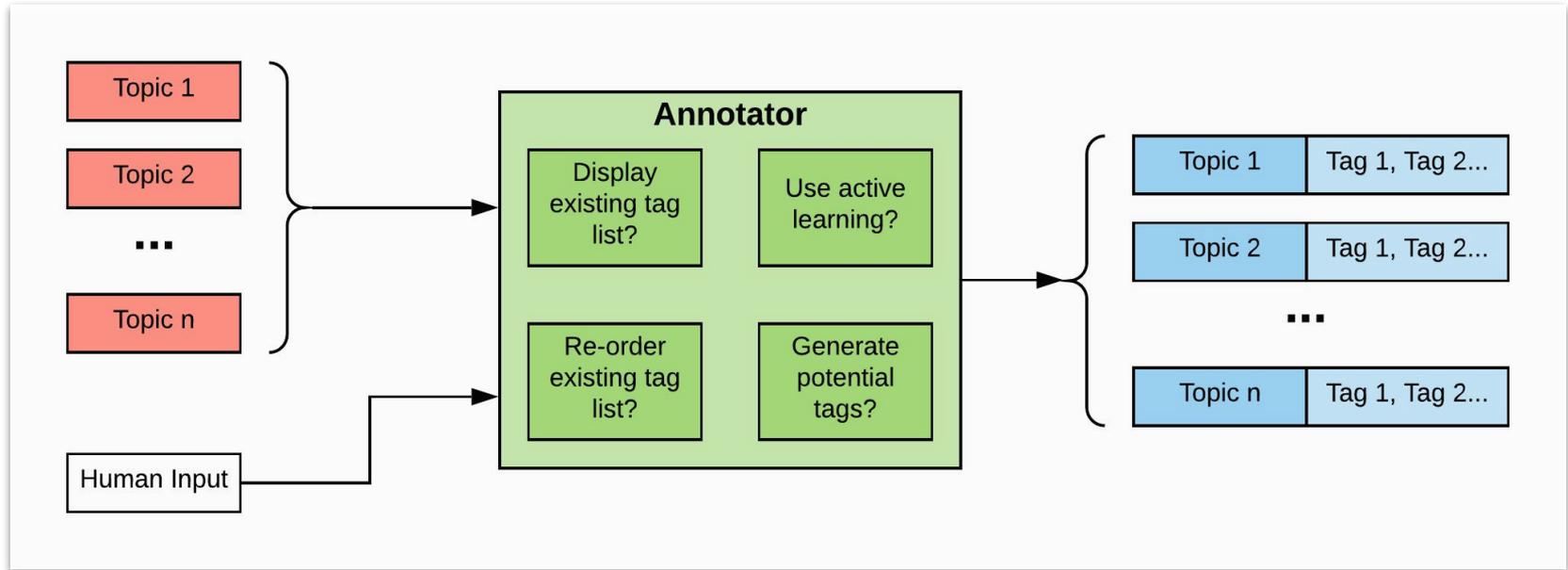
STEM-Away Tagger

Create a system to facilitate the tagging process in STEM-Away when a new topic is created



Topic Annotator

Create a system to assist in the manual annotation (labeling) of Discourse forum topics using tags



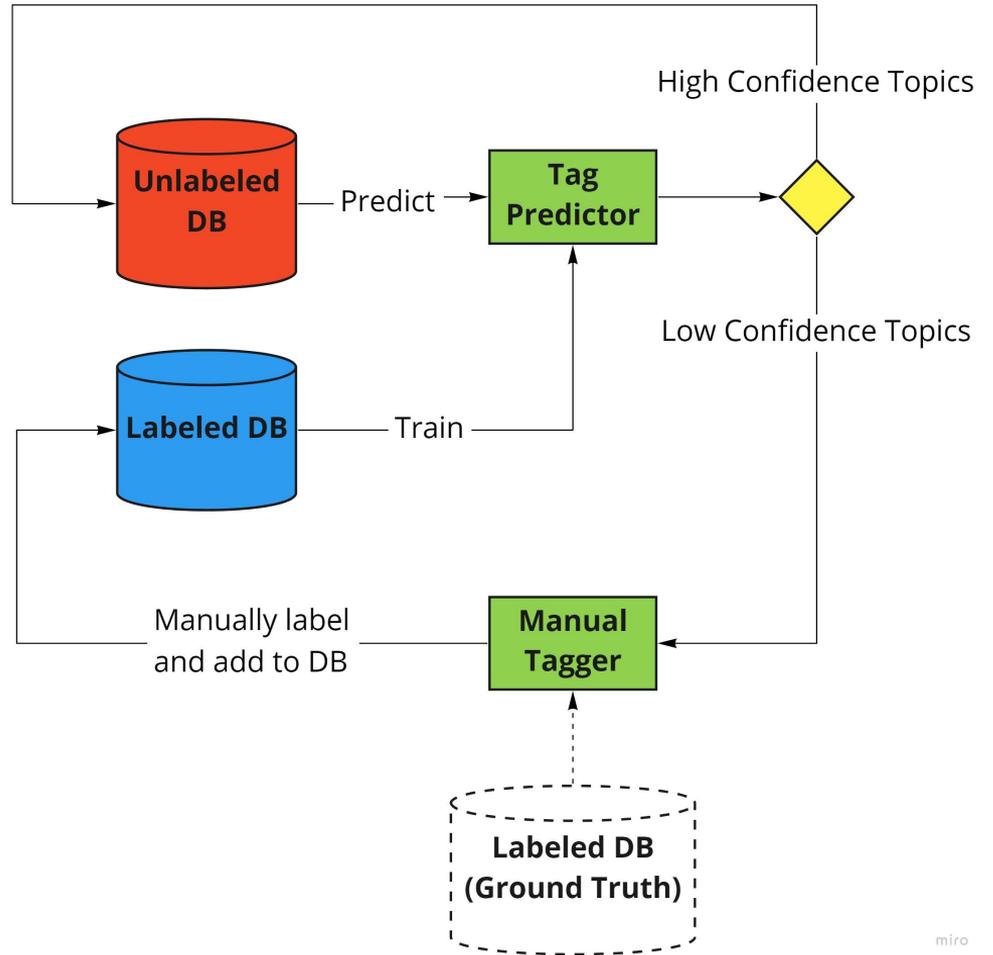
Project Goals

1. Create a modular annotation tool that uses active learning
2. Perform benchmarking analysis with Stack Exchange data
3. Deploy the annotation tool on AWS with a user-friendly interface

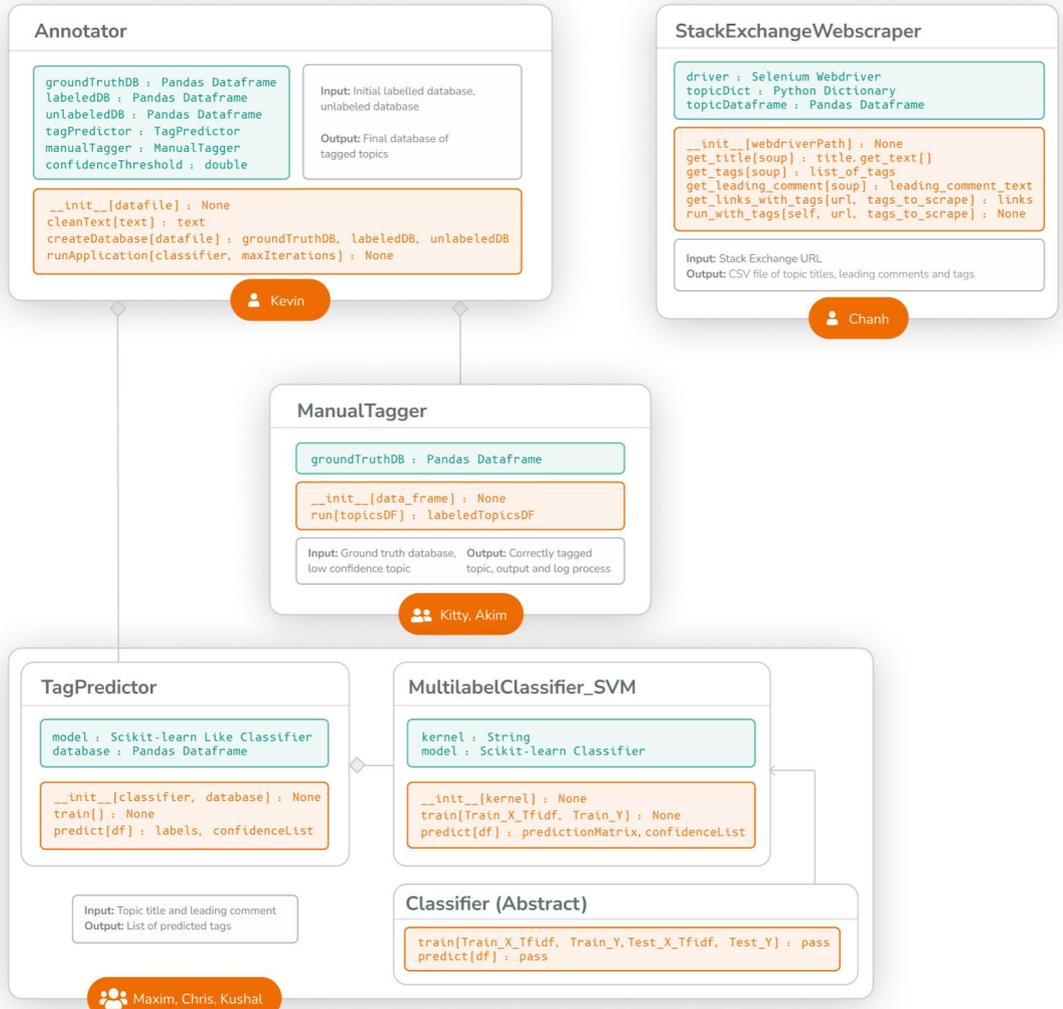
Software Architecture



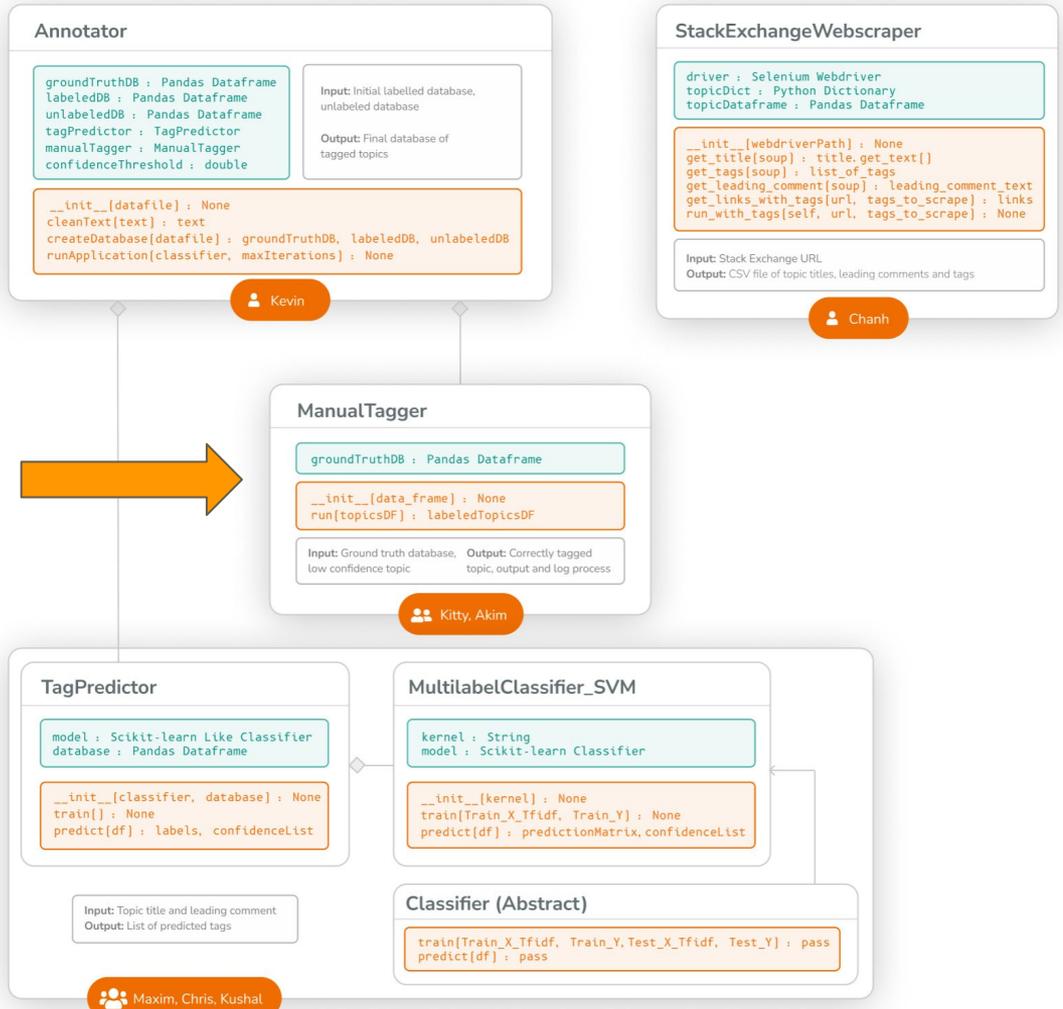
Annotator Architecture



Annotator UML Diagram



Annotator UML Diagram



Streamlit

ML July Team1 Manual Tagging App

Topic Title:

unable to get polarity scores from Vader Sentiment Anlyzer

Leading Comment:

I am trying to add these new words and their corresponding polarity scores from a csv file into Vader Sentiment LexiconIt also reflects in the vadersentiment object when it is updated:But as soon as I try to get the polarity scores for the newly added words, it throws an error:I am confused as to what is happening even though that the word is present in the vader dictionary:Does anyone know why is it happening?

Please select suitable tags for the above topic.

Choose an option

You have selected:

▶ []

IMPORTANT: Once you hit "Submit" the label is permanent. Please make sure your decision is final before hitting submit! You have 2 minutes to complete the current topic tagging

Submit

Next Topic

Dropdown menu for users to choose from

- scikit-learn
- sentiment-analysis
- web-scraping
- splinter
- text-classification
- nlk
- beautifulsoup
- nlp
- word-embedding
- text-mining
- scrapy
- selenium-webdriver
- selenium
- tf-idf

After Tagging

ML July Team1 Manual Tagging App

Topic Title:

unable to get polarity scores from Vader Sentiment Anlyzer

Leading Comment:

I am trying to add these new words and their corresponding polarity scores from a csv file into Vader Sentiment LexiconIt also reflects in the vadersentiment object when it is updated:But as soon as I try to get the polarity scores for the newly added words, it throws an error:I am confused as to what is happening even though that the word is present in the vader dictionary:Does anyone know why is it happening?

Please select suitable tags for the above topic.

sentiment-analysis x nlp x

You have selected:

▶ []

IMPORTANT: Once you hit "Submit" the label is permanent. Please make sure your decision is final before hitting submit! You have 2 minutes to complete the current topic tagging

Submit

Next Topic

Click to tag the next low confidence topic.

Hyperparameters, Metrics and Benchmarking Results

Prediction Evaluation Metrics

Accuracy:

- $(TP + TN) / (TP + TN + FP + FN)$
- Applied at the tag list level
- Counts a prediction as a failure even if one of the tags do not match the ground truth

Hamming Loss:

- $\text{sum}(\text{XOR}(\text{True}, \text{Pred})) / (\text{total number of unique tags})$
- The accuracy equation applied at the tag level
- Formally, it is the complement of tag level accuracy

Prediction Evaluation Metrics

Precision (Tag Level):

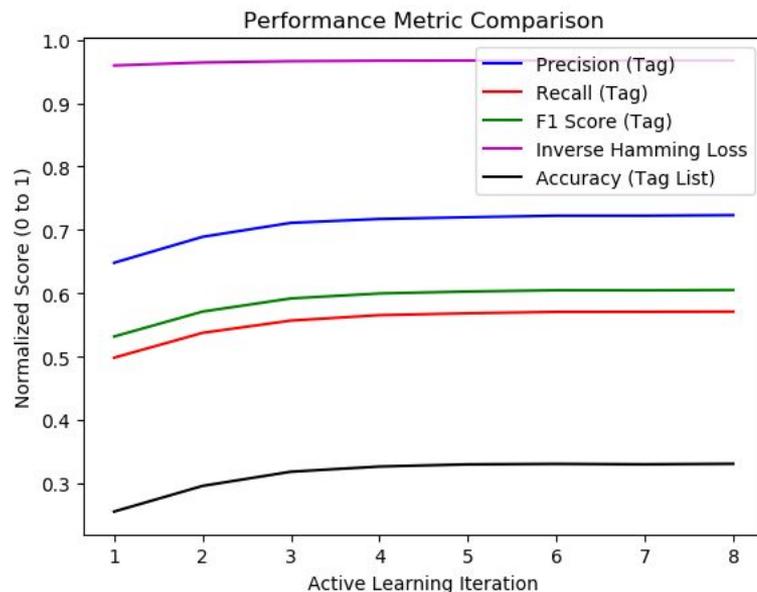
- $TP / (TP + FP)$
- Penalizes tags that should not have been flagged in the prediction

Recall (Tag Level):

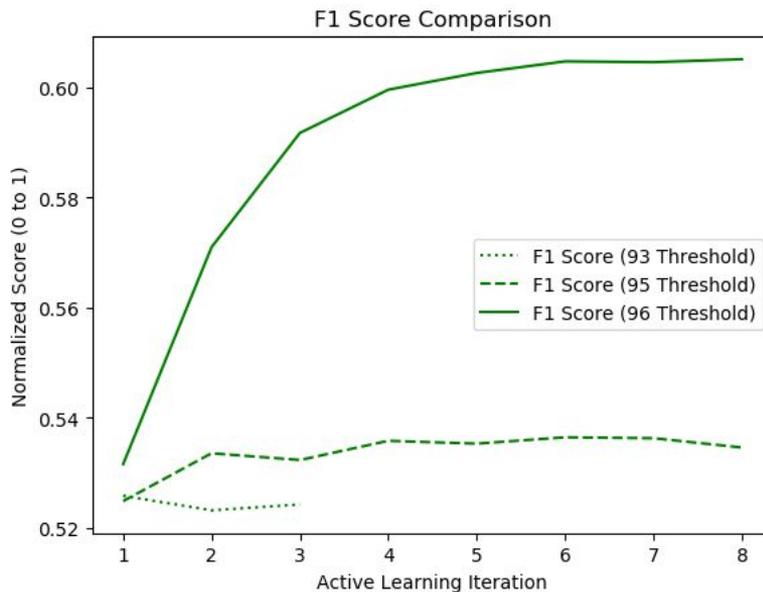
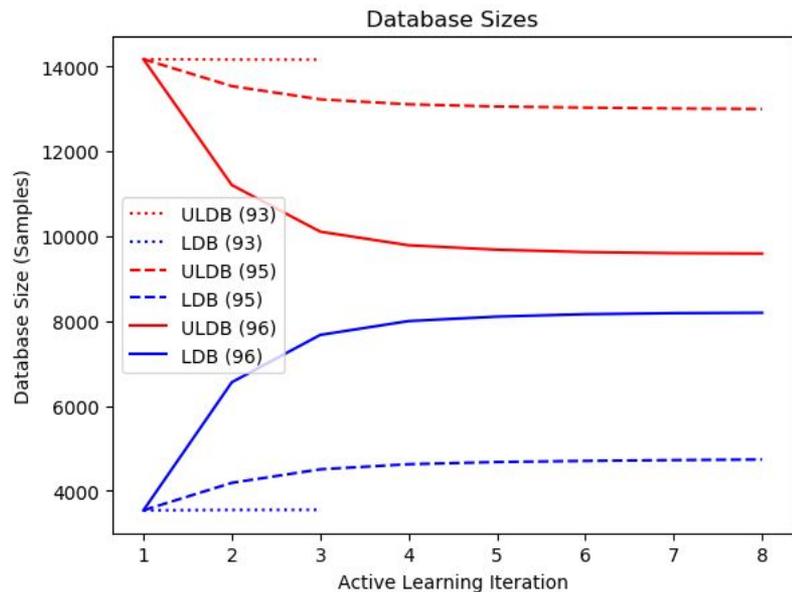
- $TP / (TP + FN)$
- Penalizes tags that should have been flagged in the prediction, but were not

F1 Score (Tag Level):

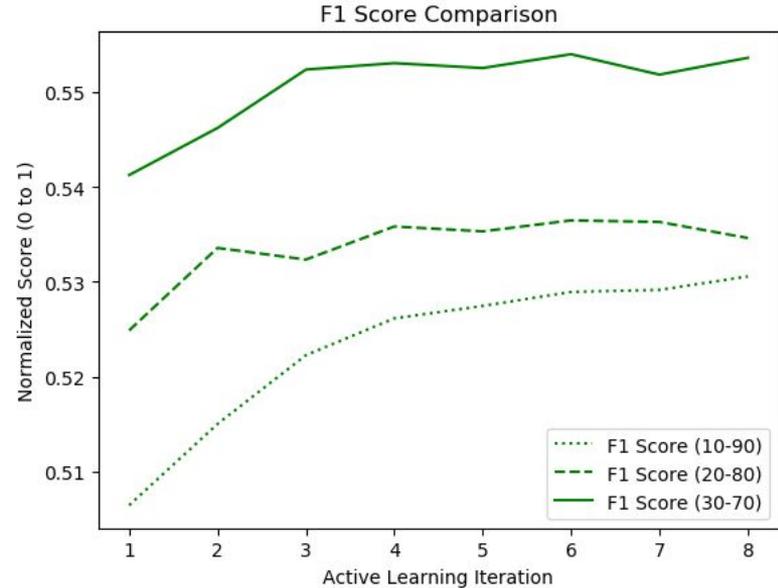
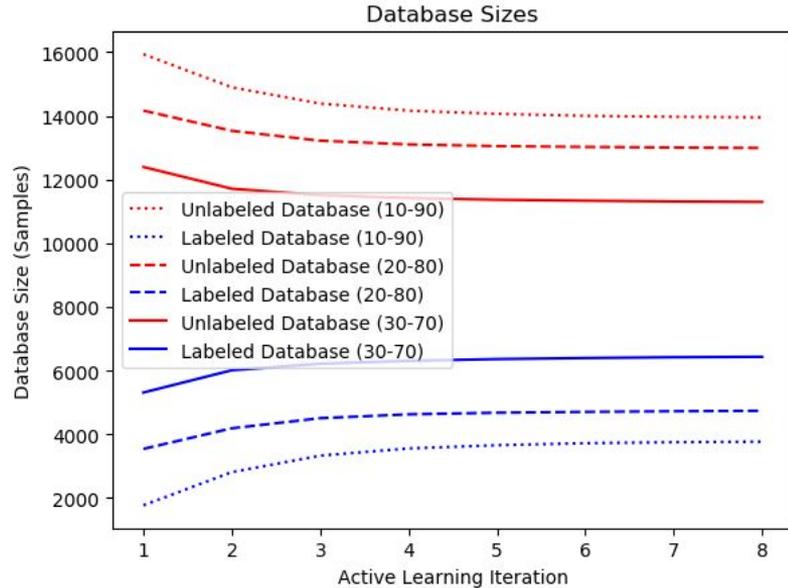
- $(2 * Precision * Recall) / (Precision + Recall)$
- Balance between both precision and recall



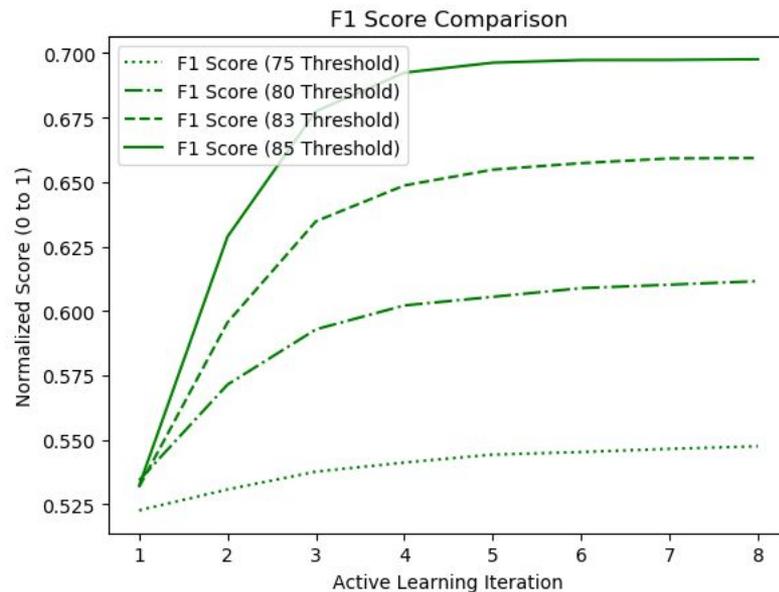
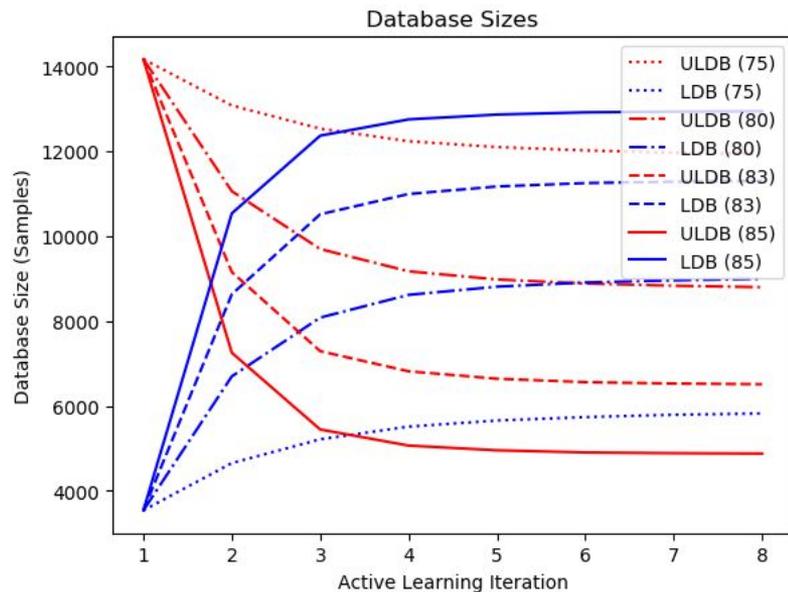
Threshold Comparison (Average Confidence, 20-80)



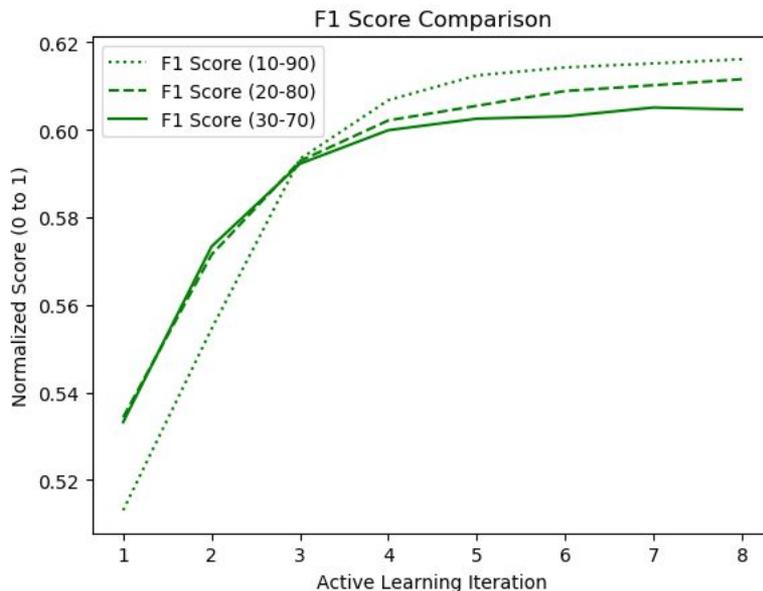
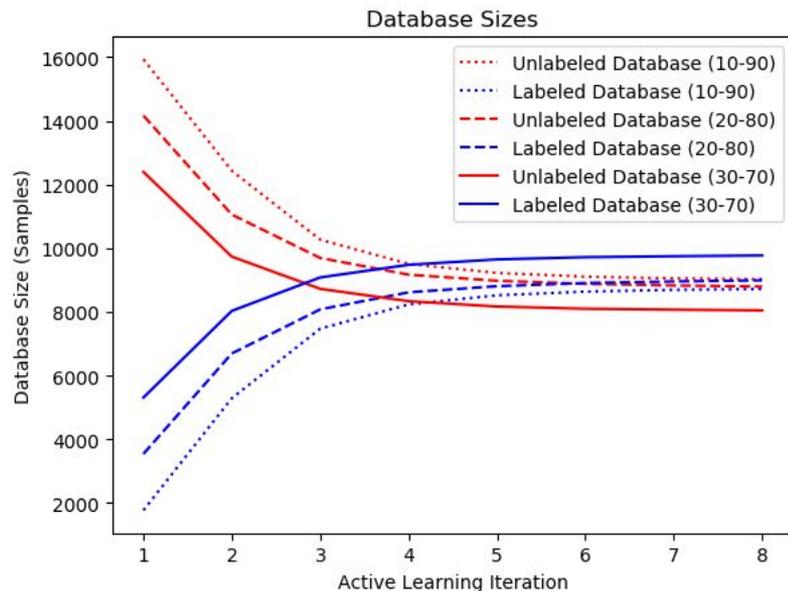
Database Size Comparison (Average Confidence, 95)



Threshold Comparison (Low-5 Confidence, 20-80)



Database Size Comparison (Low-5 Confidence, 80)



Key Conclusions

1. The Low-5 method for calculating prediction confidence is superior to the Average method, both in terms of topic selection and ease of tuning.
2. Proper tuning of the confidence threshold is important. Very low/high thresholds introduce various consequences.
3. A smaller size for the initial labeled database may result in better performance in the long run, since a larger portion of it is comprised of topics selected to optimize training, rather than selected randomly.
4. Using active learning is advantageous to not using it. A model trained on a random 90% of the full dataset (~16 000 samples) only achieved an F1 score of 56%!

Results driven but learning focused.



Collaborate: Overcame Deployment problems with great collaboration, coding happy hours and efficient self assessments.

Innovate: Brought definition to an open ended project by creative thinking and a focus on project structure

Lead: The only way to lead a team of young, smart and ambitious engineers is to lead by example and become a facilitator rather than a punisher.