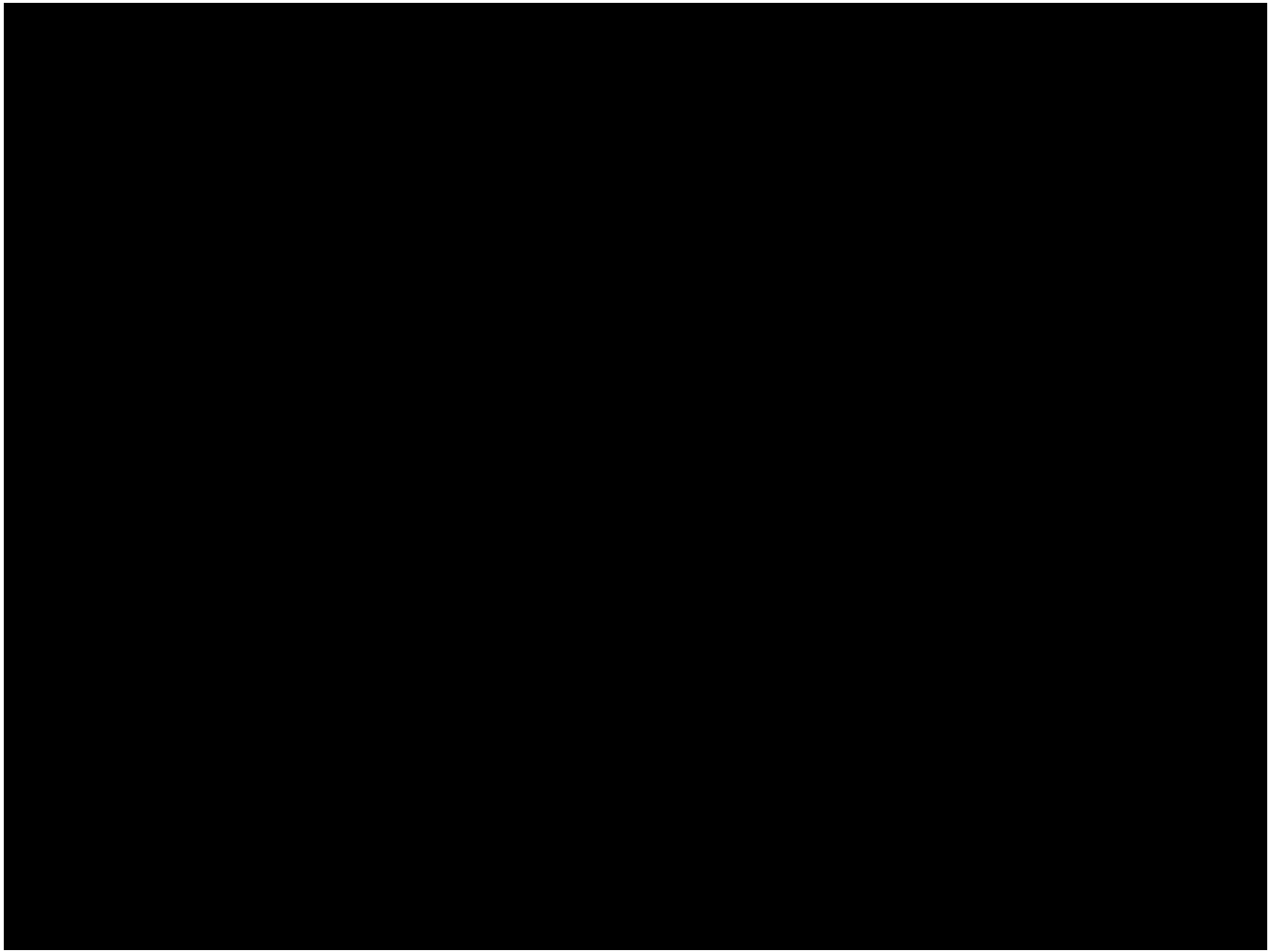
A Multi-modal Digital Assistant



Presentation structure

Chatty's GUI

- **NLP part**:

- Context Free Grammar
- Main Challenges
- Auto correct
- Heuristic based approach
- Data Augmentation
- Machine learning based approaches

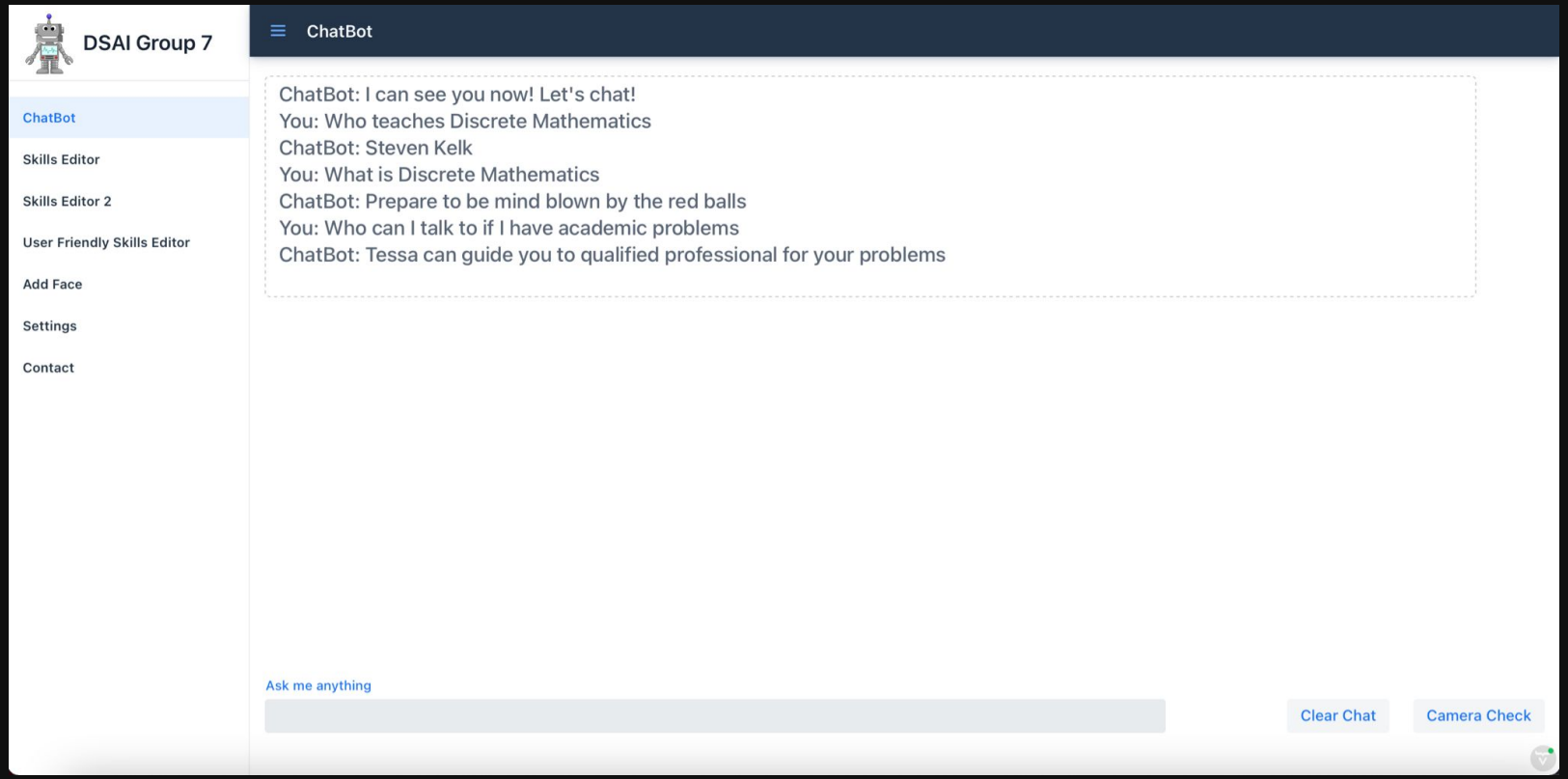
- **IVP part:**

- Skin color detection
- Haar Cascade Classifier
- Principal component analysis (PCA)

Experiment and results

Parser

Vaadin


DSAI Group 7
ChatBot
Skills Editor
Skills Editor 2
User Friendly Skills Editor
Add Face
Settings
Contact
ChatBot
ChatBot: I can see you now! Let's chat!
You: Who teaches Discrete Mathematics
ChatBot: Steven Kelk
You: What is Discrete Mathematics
ChatBot: Prepare to be mind blown by the red balls
You: Who can I talk to if I have academic problems
ChatBot: Tessa can guide you to qualified professional for your problems
Ask me anything
Clear Chat
Camera Check

GUI

CFG-based skills

Variable:	Expression:
<S>	<ACTION>
<ACTION>	<FOOD> | <SCHEDULE>
<FOOD>	Where can I eat <DISH>
<DISH>	sushi
...	

- Too few skills...
- Encode up to 52 skills !

Parser

Main challenges

- Lack of Data
- Deal with synonyms
- Spelling mistakes

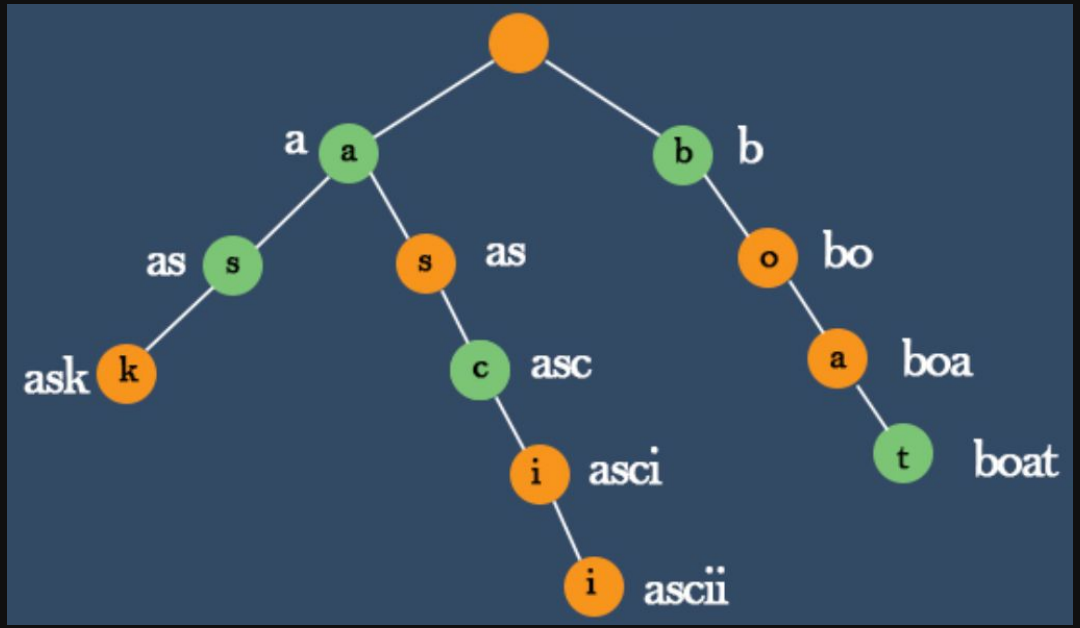
Benchmark Autocorrect

- Levenshtein distance
- Trie Data structure
- Edit distance

Benchmark
Autocorrect

$$\operatorname{lev}(a,b) = \begin{cases} |a| & \text{if } |b| = 0, \\ |b| & \text{if } |a| = 0, \\ \operatorname{lev}(\operatorname{tail}(a), \operatorname{tail}(b)) & \text{if } a[0] = b[0] \\ 1 + \min \begin{cases} \operatorname{lev}(\operatorname{tail}(a), b) \\ \operatorname{lev}(a, \operatorname{tail}(b)) \\ \operatorname{lev}(\operatorname{tail}(a), \operatorname{tail}(b)) \end{cases} & \text{otherwise.} \end{cases}$$

Benchmark Autocorrect



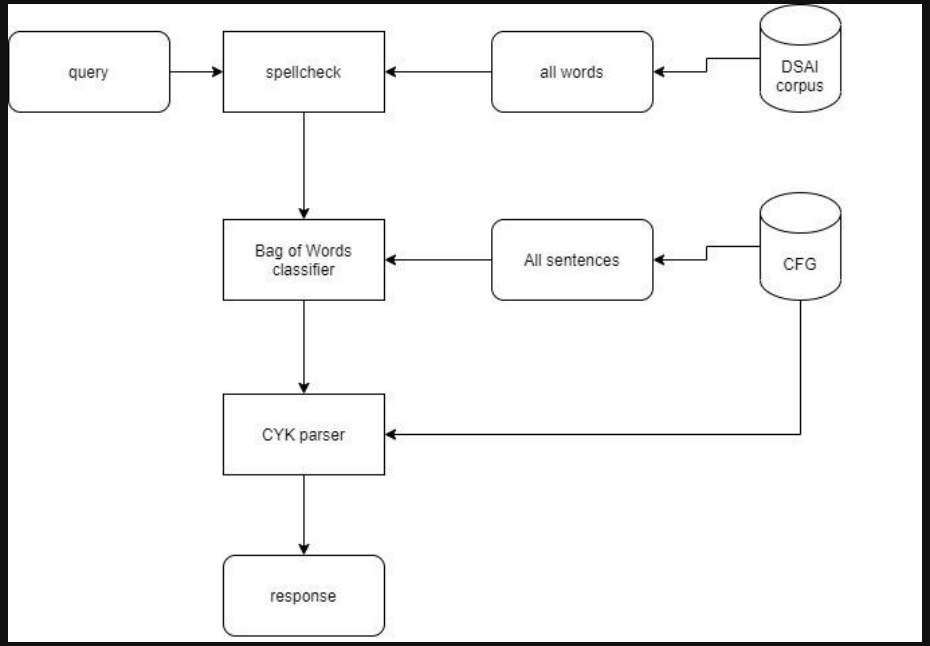
Autocorrect

Autocorrects 88% of misspelled words with 1 mutation.
Autocorrects 72.7% of misspelled words with 2 mutations.

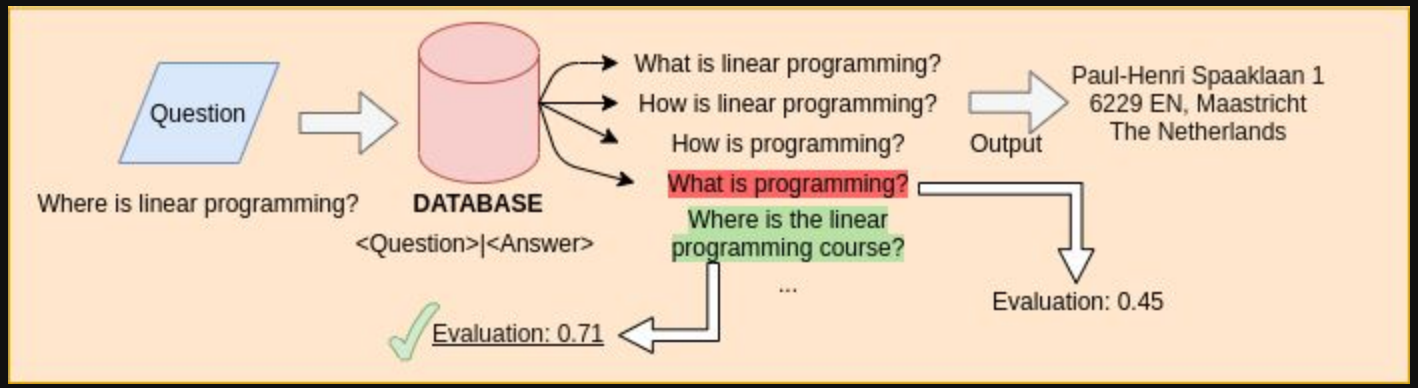
Benchmark Autocorrect

DRAWBACKS:

- **Dictionary needs to be comprehensive**
- **Not all types of misspellings are captured**
- **Context independent**



Benchmark Approach - Bag Of Words Based Chatbot



Benchmark Approach - Heuristic Based Chatbot

Evaluation Function

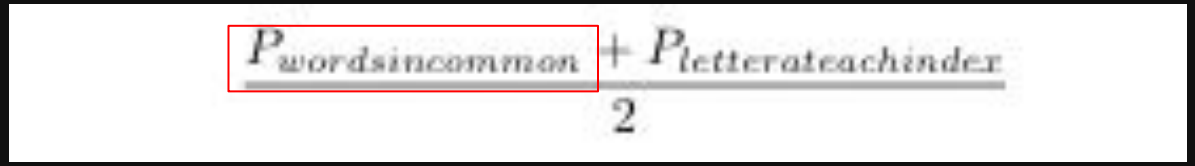
Benchmark Approach - Heuristic Based Chatbot

$$\frac{P_{wordsincommon} + P_{letterateachindex}}{2}$$

Benchmark Approach - Heuristic Based Chatbot

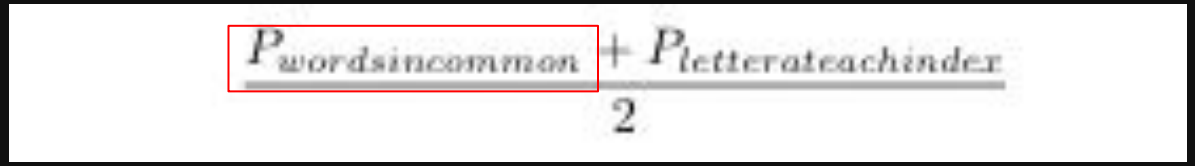
$$\frac{P_{wordsincommon} + P_{letterateachindex}}{2}$$

Benchmark Approach - Heuristic Based Chatbot



Database question: What is linear programming?

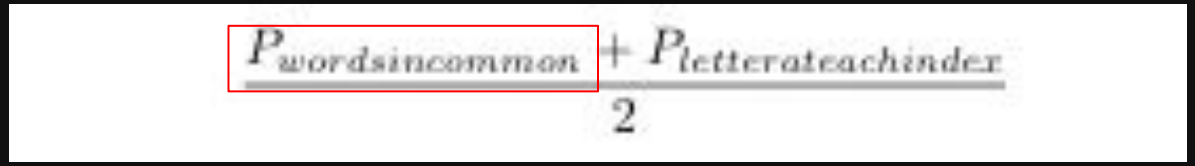
Benchmark Approach - Heuristic Based Chatbot



Database question: What is linear programming?

User Query: When is linear programming?

Benchmark Approach - Heuristic Based Chatbot



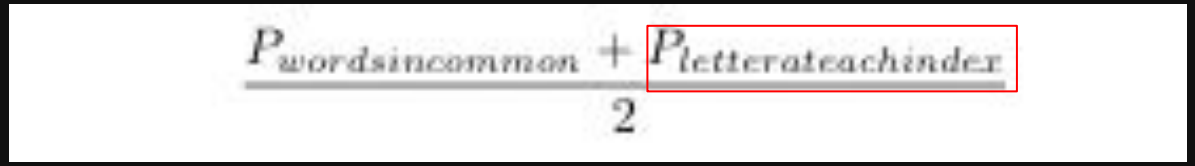
Database question: What is linear programming?

User Query: When is linear programming?

Benchmark Approach - Heuristic Based Chatbot

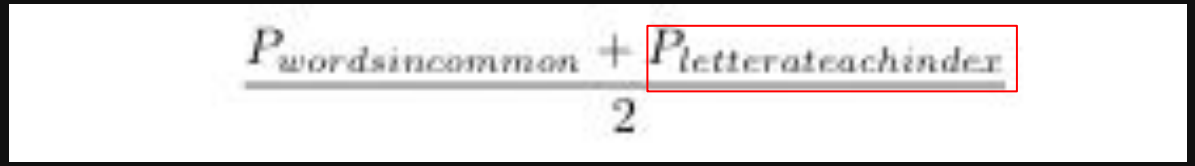
$$\frac{P_{wordsincommon} + P_{letterateachindex}}{2}$$

Benchmark Approach - Heuristic Based Chatbot



Database question: What is linear programming?

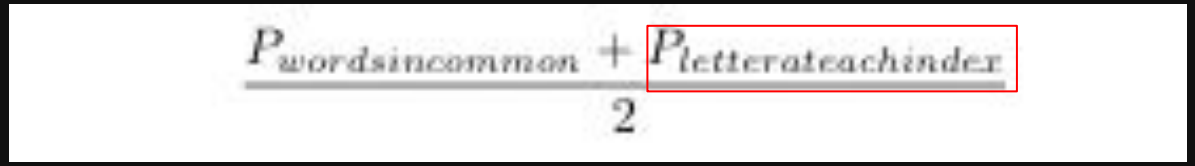
Benchmark Approach - Heuristic Based Chatbot



Database question: What is linear programming?

User Query: When is linear programming?

Benchmark Approach - Heuristic Based Chatbot



Database question: What is linear programming?

User Query: When is linear programming?

Benchmark Approach - Heuristic Based Chatbot

What if there is a typo?

$$\frac{P_{wordsincommon} + P_{letterateachindex}}{2}$$

Database question: What is linear programming?

User Query: Whay is linear programnning?

Benchmark Approach - Heuristic Based Chatbot

Issues?

Does not take into account context

Database question: When do we have linear programming

Benchmark Approach - Heuristic Based Chatbot

Database question: When do we have linear programming

User Query 1: Why do we have linear programming?

Benchmark Approach - Heuristic Based Chatbot

Database question: When do we have linear programming

User Query 1: Why do we have linear programming? 0.78

Benchmark Approach - Heuristic Based Chatbot

Database question: When do we have linear programming

User Query 1: Why do we have linear programming? 0.78

User Query 2: At what time is linear programming? 0.53

Benchmark Approach - Heuristic Based Chatbot

Database question: When do we have linear programming

User Query 1: Why do we have linear programming? 0.78

User Query 2: At what time is linear programming? 0.53

Benchmark Approach - Heuristic Based Chatbot

Database question: When do we have linear programming

User Query 1: Why do we have linear programming? 0.78

Answer: Because it is important.

Benchmark Approach - Heuristic Based Chatbot

Data Augmentation

The more data the better - Group 07

Data Augmentation

Issues?

No data at all.

Solution:

Data Augmentation

Solution: generate some data and then augment it

Data Augmentation

Backtranslation?

Backtranslation
impossible because of too short sentences

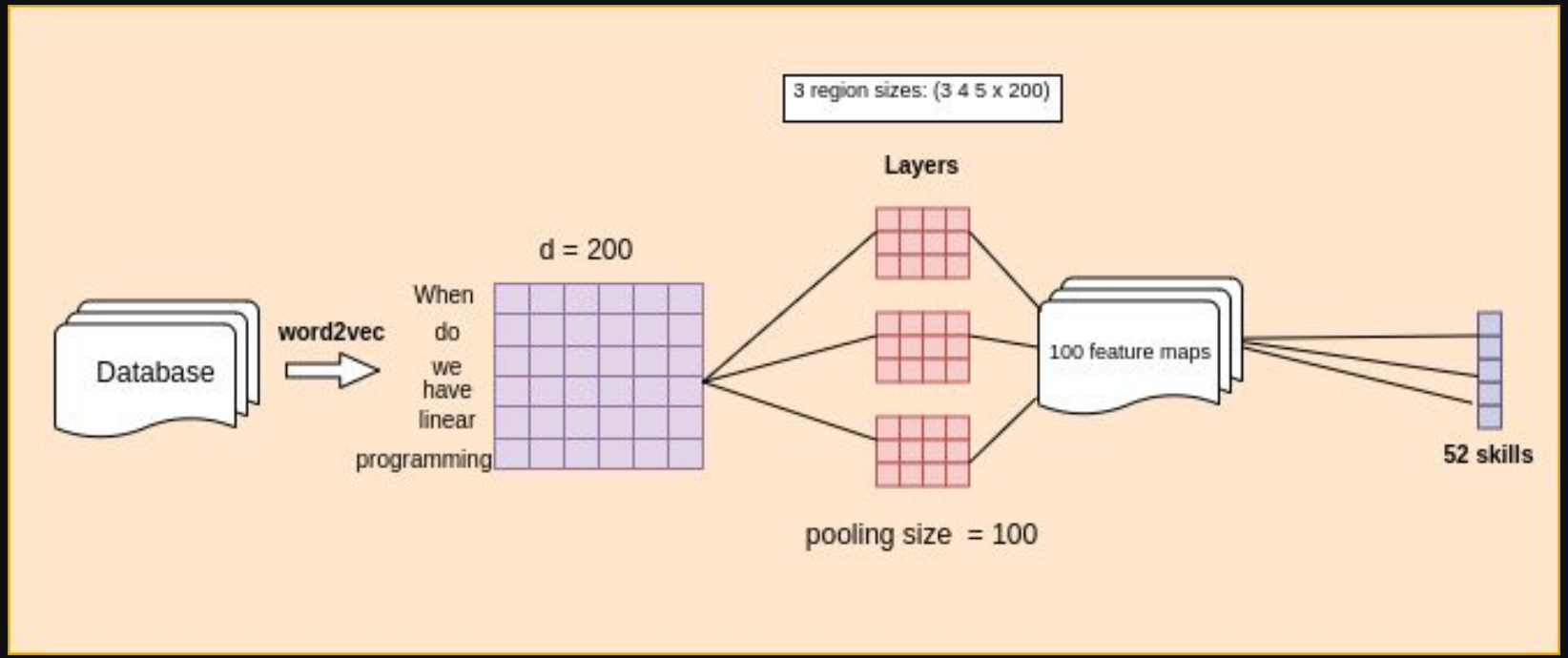
Augment the data using rephrasing

Augment the data using rephrasing

(uses BERT under the hood)

Machine Learning based approaches

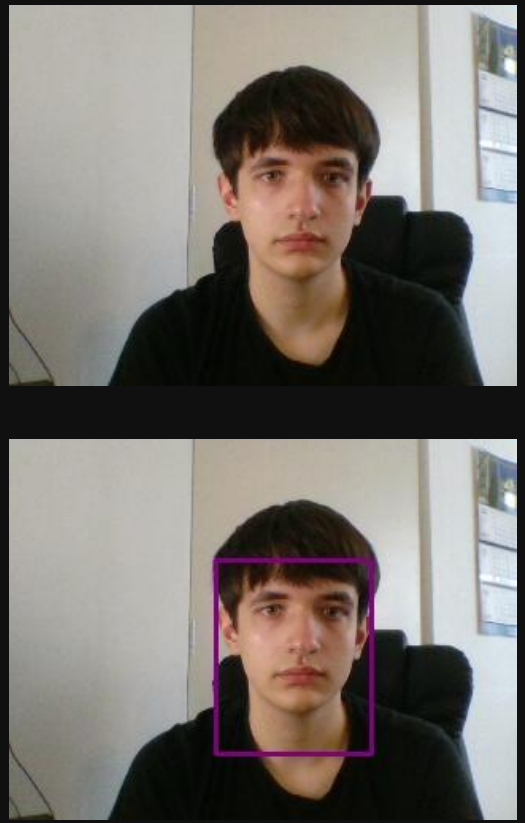
CNN structure



ML-based approach

Skin Color Detection

- Linear separation of skin and non-skin tones
- Color spaces used for recognizing skin are RGB, HSV and YCbCr
- Rule-based separation
- Easy to implement and shows good results
- After finding the mask, we extract boundaries
- In the final stage we only keep the boundaries that are the size of a face and draw a bounding box around those boundaries
- Morphological opening to pull apart connected regions



Face Detection

RGB Mask

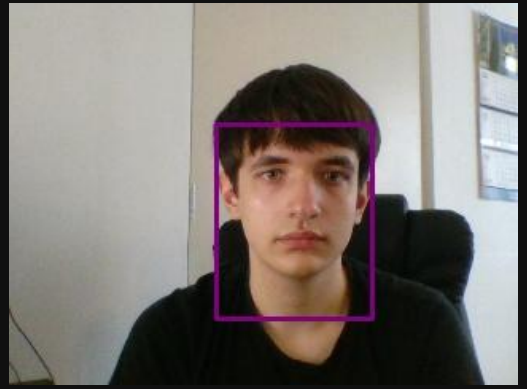
YCbCr Mask

HSV Mask

Combined Mask

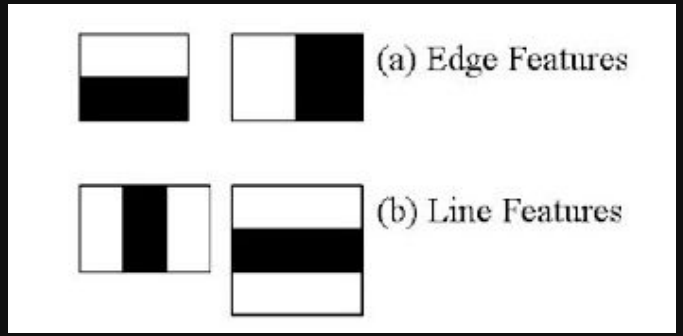
Contour Extraction

Bounding Box

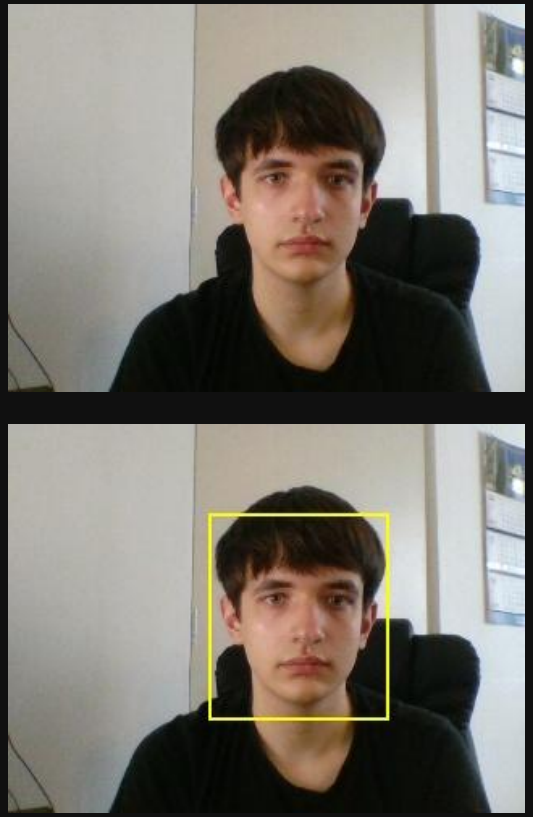


HAAR CLASSIFIERS

- Haar features are extracted from lots of positive images (containing faces)
- The features work like convolutional kernels



- Since there are lots of features to calculate, the process is sped up by using integral images
- Best features are then selected by **Adaboost training**
- The concept of Cascade of Classifiers is introduced. Instead of applying all features at once they are grouped and applied one-by-one.



Face Detection

Principal Component Analysis (PCA)

PROBLEM:

Viewing each picture as a vector

SOLUTION:

Extract most important dimensions (main features of the face)

How it works ?

- Compute covariance matrix from the initial training set of face images
- Obtain the eigenvalues of the covariance matrix as well as the eigenvectors

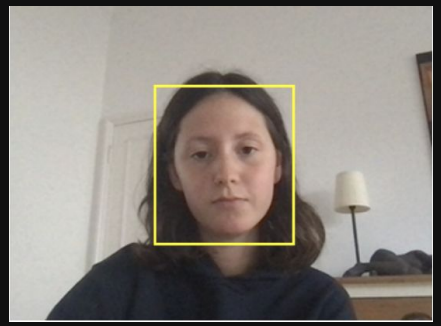
Training phase

Training set of faces → Compute eigenfaces → Eigenspace

- Lower dimension
- Contains main components

Recognition phase

HCC



Project input image into each of the eigenface of the eigenspace

Select the minimum distance

If lower than threshold >> label with the face of the eigenface giving the minimum distance

Experiments & Results

1) **Is it possible to implement a skin color model that is going to perform as well as a Haar Cascade Classifier?**

Algorithms	FDR (%)	DSR (%)
Haar Cascade Classifier	18.82	99.33
RGB-HSV-YCbCr	**20.26**	**81.33**
RGB-HSV	26.72	73.55
RGB	37.84	64.22

2) What is the optimal number of eigenvectors to successfully recognize faces using PCA (Principal Component Analysis) with satisfactory accuracy and computational speed?

Eigenvectors Count	DSR (%)	Total Time (s)
10	36	66.2
20	42	70
40	46	71.5
80	**52**	**76.9**
160	52	116.9
All	52	212

3) What is the highest accuracy that can be reached for recognizing utterances from the CFG, and what model achieves this accuracy?

Number of class	2
Accuracy	0.8283
Precision	0.8367
Recall	0.9462
F1 Score	0.8881

Bag of Words Approach

Number of class	2
Accuracy	0.6454
Precision	0.6308
Recall	0.8367
F1 Score	0.7193

CNN

NLP

Trieu Ngoc Hoang
Grassetto Nico
De Clercq William

VISION

Roşca Alexei
Imparato Adele

Thank you!