FEDERICO MINNITI, MATTEO DEL SEPPIA

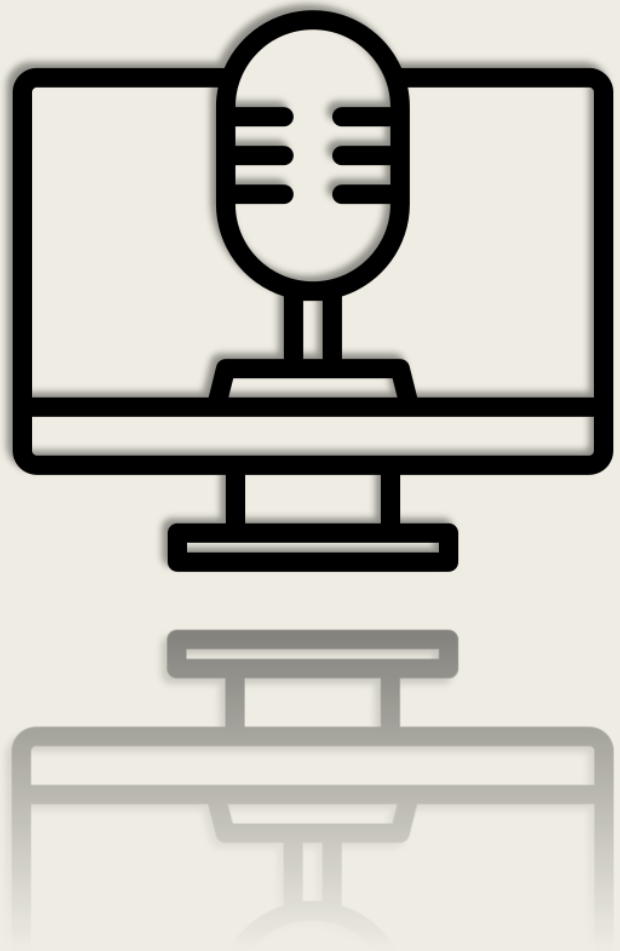
DATA MINING AND MACHINE LEARNING

HIGHLIGHTS

- VoiceIDNotes is a private note taking application that allows users to create, store, and manage their notes

- The key idea is that everyone has their own particular and personal voice. VoiceIDNotes allows users to have an easy login phase because once recognized their voices a simple 4-digit pin will be required

- Users record a small voice audio and from that, audio features are extracted: mfccs, delta and deltadelta.

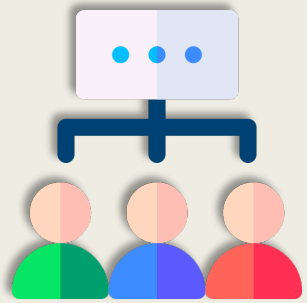


DATASET DESCRIPTION

- Source: https://www.openslr.org/12
- Description: LibriSpeech is a corpus of approximately 500 hours of 16kHz read English speech. The data is derived from read audiobooks from the LibriVox project.
- Volume: 30 GB with 500 hours of speech with lots of different speakers
- From the whole list of speakers we selected 40, including 20 men and 20 women

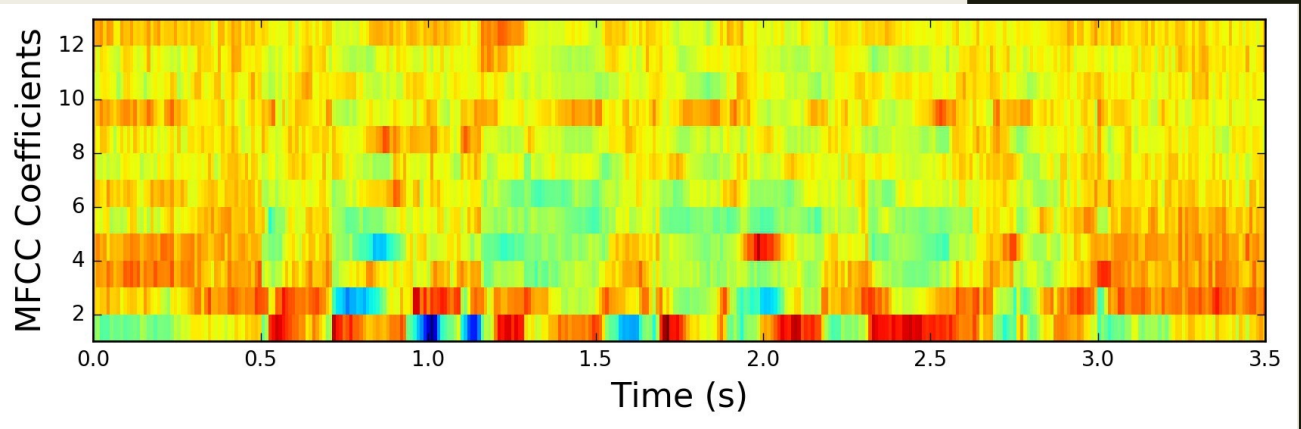
VOICE AUDIO
PRE-PROCESSING

LEARNING



CLASSIFICATION

MACHINE LEARNING



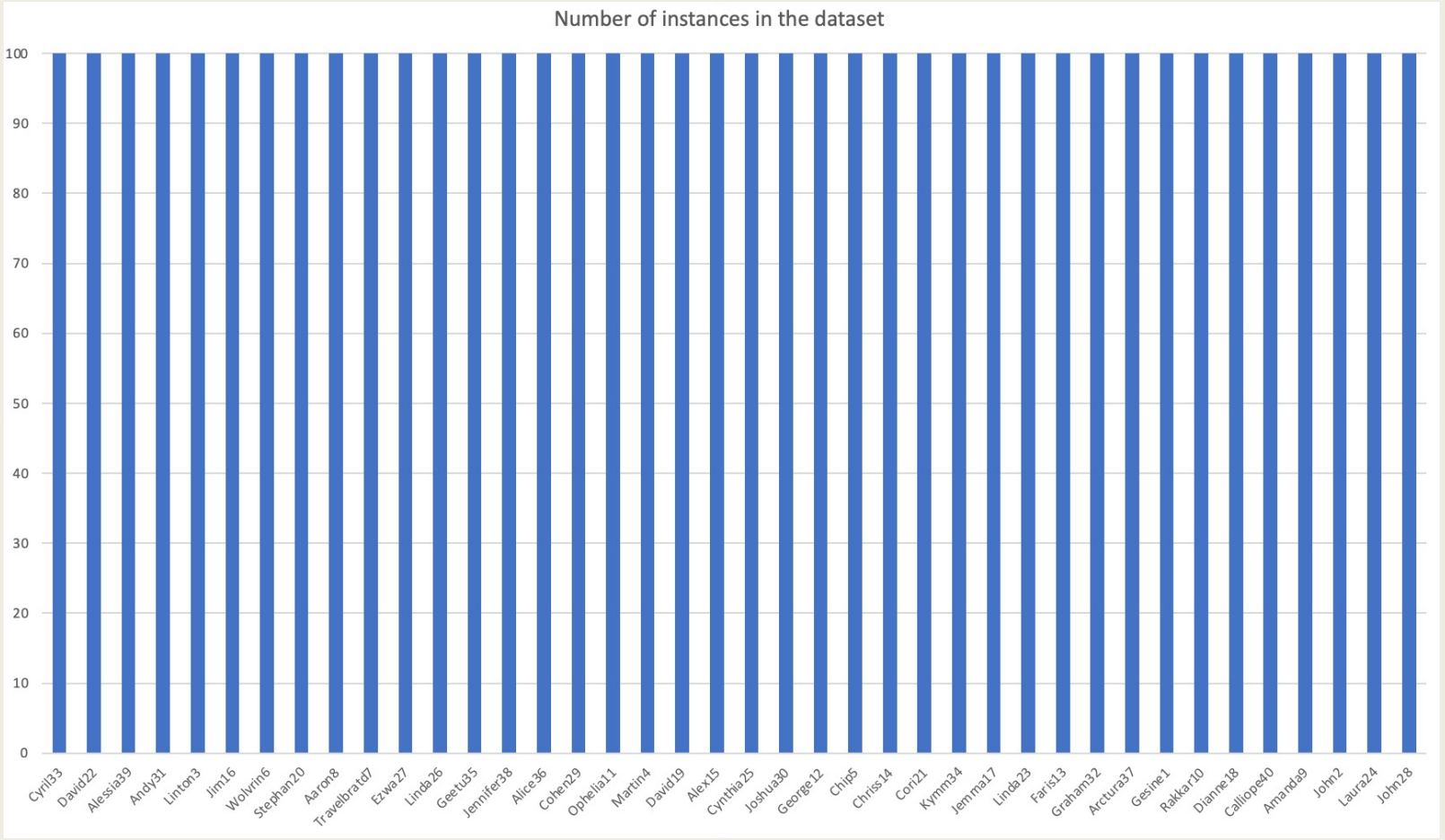
AUDIO FEATURES EXTRACTION

13 MFCCs: they are a small set of features which concisely describe the overall shape of the spectrum of the amplitude of the signal

13 Deltas: calculated as variations of the MFCCs.

13 Delta-Deltas: extracted as the variation of the Delta coefficients

DATA TRANSFORMATION



Normalization: z-score

MODELS EVALUATION (1/3)

Algorithm	Attribute Selection	# Selected Attributes	Accuracy	Avg Precision	Avg Recall	Avg F-measure	Tree dimension	Time to build model
J48	null	1-39 (tot. 39)	88,7	0,889	0,887	0,888	397	0.22s
J48	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot. 14)	88,5	0,888	0,886	0,886	407	0.27s
J48	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	88,6	0,886	0,886	0,884	409	0.12s
Naive Bayes	InfoGainAttributeEval + Ranking (0.9)	1-13 (tot. 13)	93,6	0,938	0,936	0,936		0.11s
Naive Bayes	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	93,7	0,94	0,938	0,938		0.03s
Naive Bayes	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot. 14)	93,6	0,939	0,936	0,936		0.16s
RandomForest	null	1-39 (tot. 39)	98,5	0,985	0,985	0,985		3.10s
RandomForest	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot. 14)	98,3	0,984	0,984	0,983		2.21s
RandomForest	InfoGainAttributeEval + Ranking (0.9)	1-13 (tot .13)	98,45	0,985	0,985	0,984		2.09s

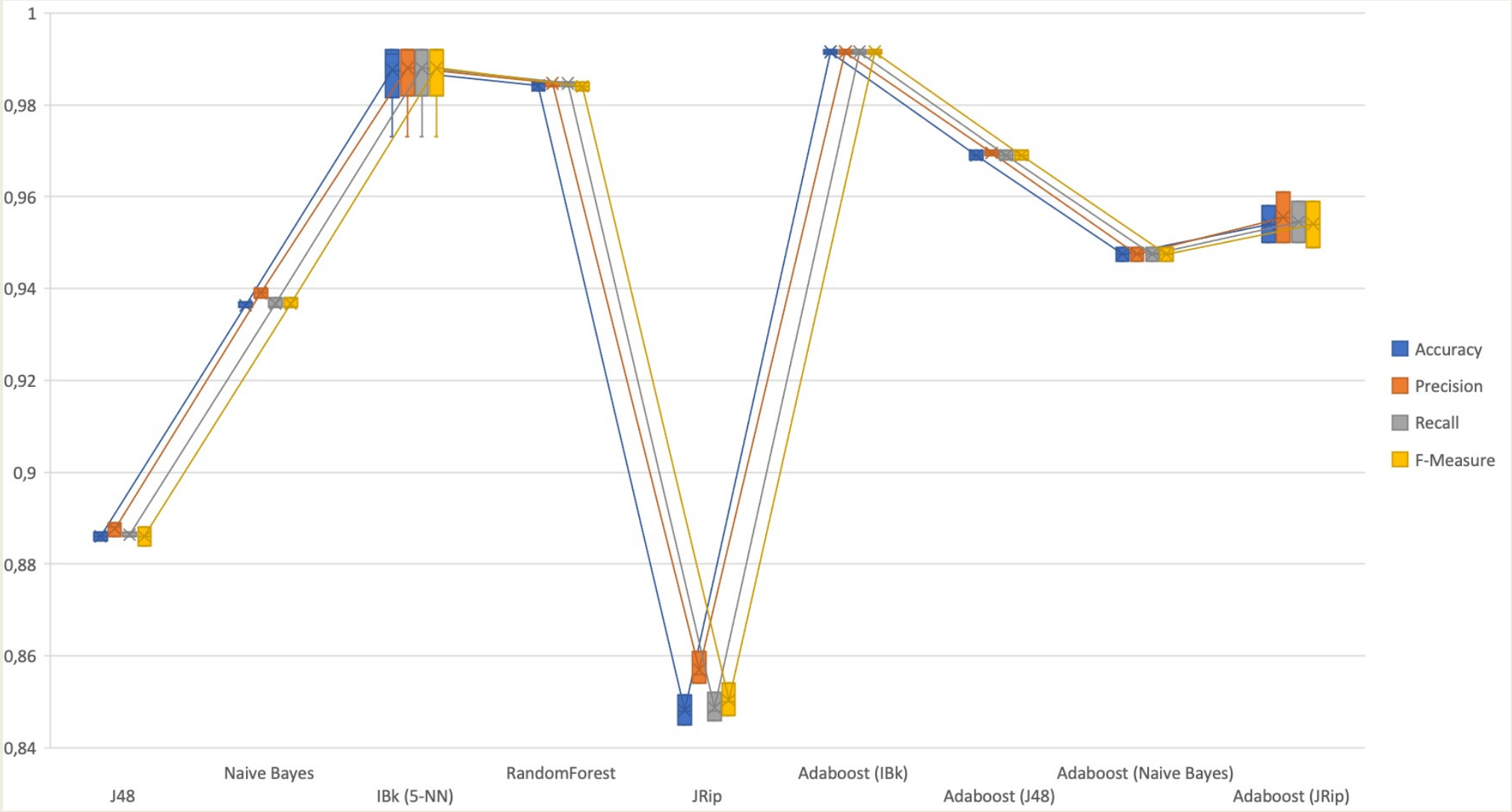
MODELS EVALUATION (2/3)

Algorithm	Attribute Selection	# Selected Attributes	Accuracy	Avg Precision	Avg Recall	Avg F-measure	Tree dimension	Time to build model
JRip	CFSubsetEval + GreedyStepWise	1-13, 27 (tot. 14)	84,8	0,854	0,848	0,85		1.45s
JRip	InfoGainAttributeEval + Ranking (0.9)	1-13 (tot. 13)	85,15	0,861	0,852	0,854		1.56s
JRip	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	84,5	0,856	0,846	0,847		1.36s
IBk (5-NN)	null	1-39 (tot. 39)	97,3	0,973	0,973	0,973		0.01s
IBk (5-NN)	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot. 14)	99,2	0,992	0,992	0,992		0.14s
IBk (5-NN)	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	99,1	0,992	0,992	0,992		0.03s
IBk (5-NN)	InfoGainAttributeEval + Ranking (0.9)	1-13 (tot. 13)	99	0,991	0,991	0,991		0.1s
IBk (5-NN)	CFSubsetEval + GreedyStepWise	1-13, 27 (tot. 14)	99,2	0,992	0,992	0,992		0.14s

MODELS EVALUATION (3/3)

Algorithm	Attribute Selection	# Selected Attributes	Accuracy	Avg Precision	Avg Recall	Avg F-measure	Tree dimension	Time to build model
Adaboost (IBk)	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	99,1	0,991	0,991	0,991		9.41s
Adaboost (IBk)	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot.14)	99,2	0,992	0,992	0,992		9.81s
Adaboost (J48)	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot.13)	97	0,97	0,97	0,97		1.91s
Adaboost (J48)	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot.14)	96,8	0,969	0,968	0,968		2.13s
Adaboost (Naive Bayes)	CorrelationAttributeEval + Ranking(0.1)	1-13 (tot. 13)	94,6	0,946	0,946	0,946		2.19s
Adaboost (Naive Bayes)	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot. 14)	94,9	0,949	0,949	0,949		2.95s
Adaboost (JRip)	CFSubsetEval + BestFirst (Backward)	1-13, 27 (tot.14)	95,8	0,961	0,959	0,959		11.64s
Adaboost (JRip)	InfoGainAttributeEval + Ranking (0.9)	1-13 (tot.13)	95	0,95	0,95	0,949		11.5s

MODELS SELECTION

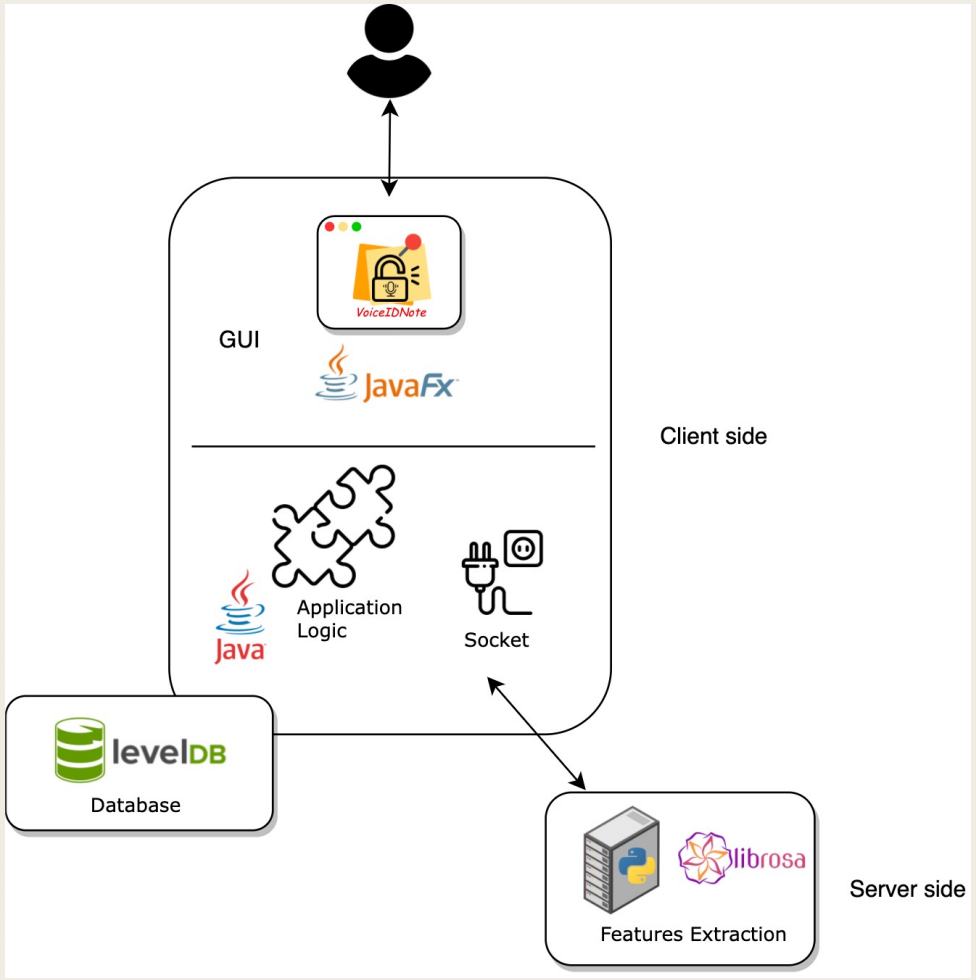


STATISTICAL SIGNIFICANCE

	IBk (k = 5) CorrelationAttributeEval + Ranking(0.1)	Naive Bayes CorrelationAttributeEval + Ranking(0.1)	Random Forest InfoGainAttributeEval + Ranking (0.9)	Adaboost (J48) CorrelationAttributeEval + Ranking(0.1)
Accuracy	99,16	93,93 *	98,52 *	96.96*
F-measure	0.99	0.99	0.99	0.99
Time for testing	0.07s	0.02s *	0.02s *	0.01s *

THE APPLICATION

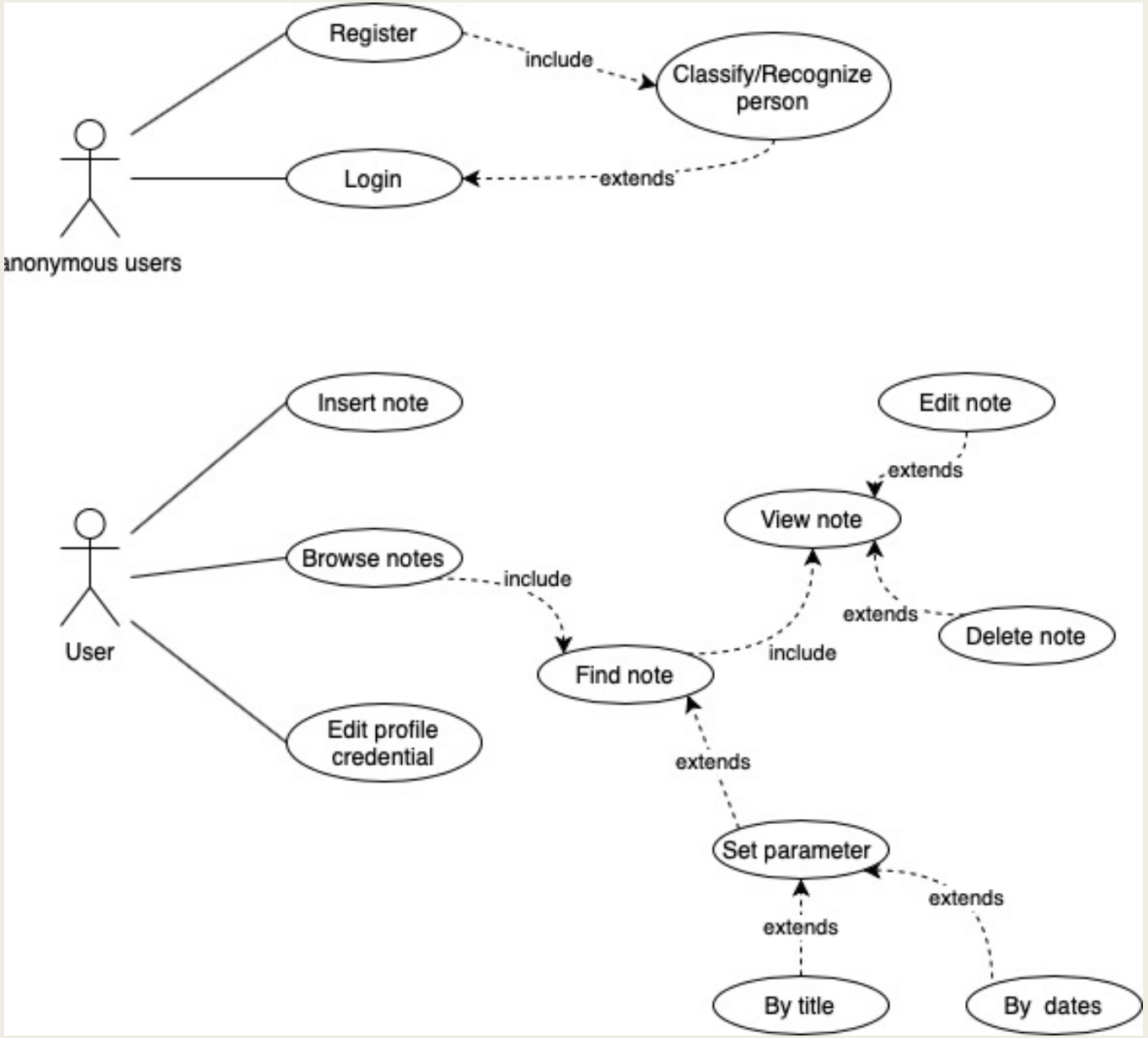
SYSTEM ARCHITECTURE



Architectural pattern: client-server
Client divided in:
-> GUI
-> Application logic
-> Database

Server:
-> Features extraction server

USE CASE DIAGRAM



CLASS DIAGRAM

DATA MODEL

Key schema:

- User
  - user:<usename>:password
  - user:<usename>:pin
- Note
  - note:<username>:<timestamp>:title
  - note:<username>:<timestamp>:text

LOGIN

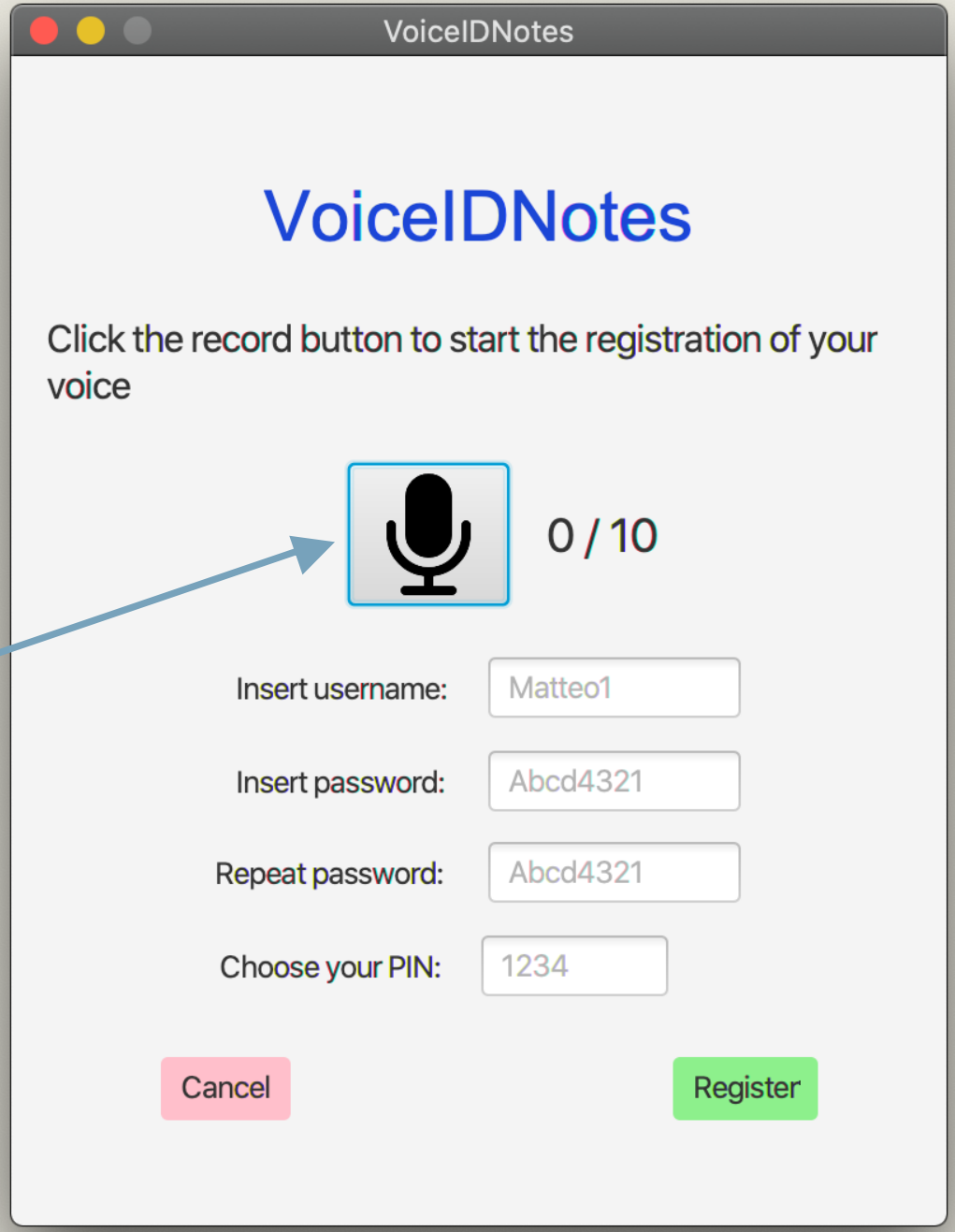
If it correctly recognise you, put your PIN

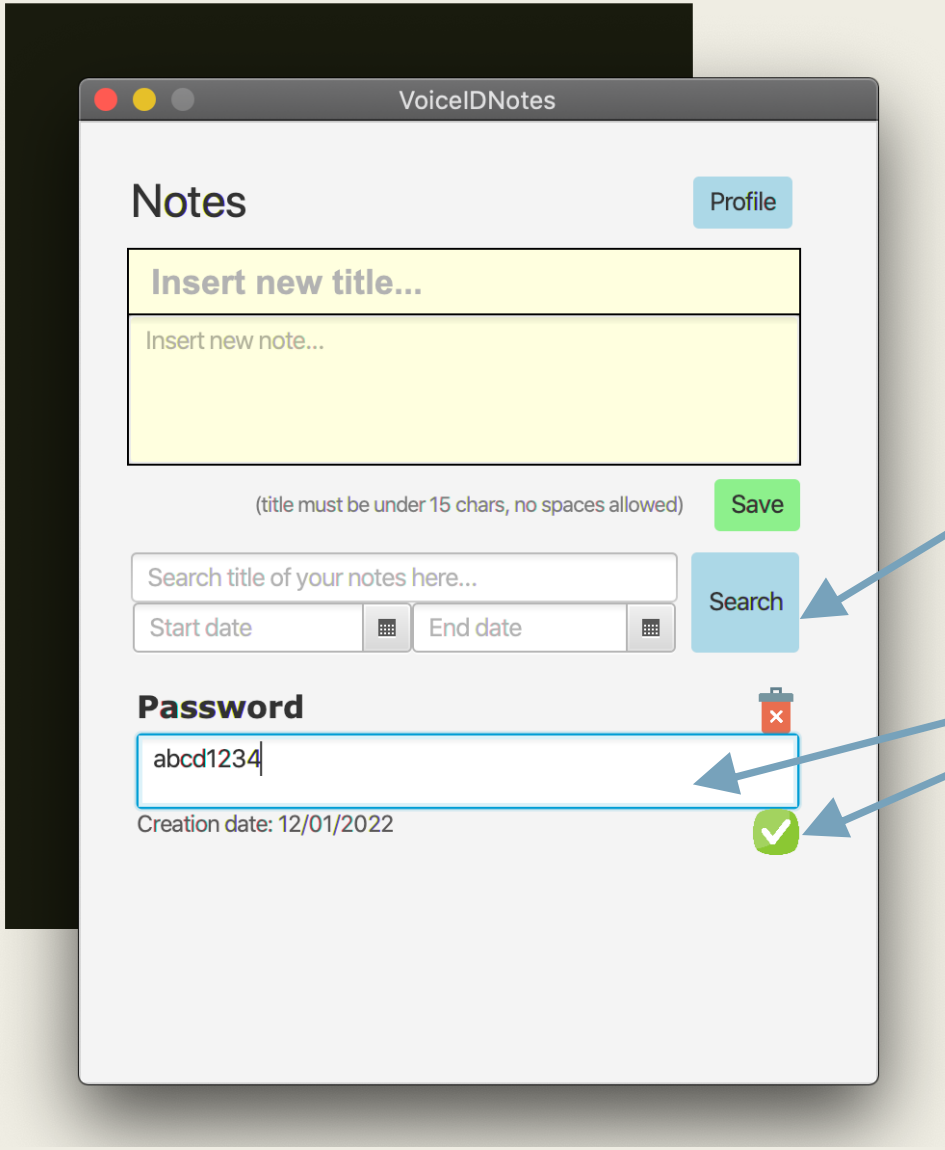
Click Register button to create an account

Click login button to access

REGISTER

Click to start recording 10 audio
then some example sentences appear





Filter and search notes

Click in the text area to modify a note and then the update button appears

NOTES

PROFILE PAGE