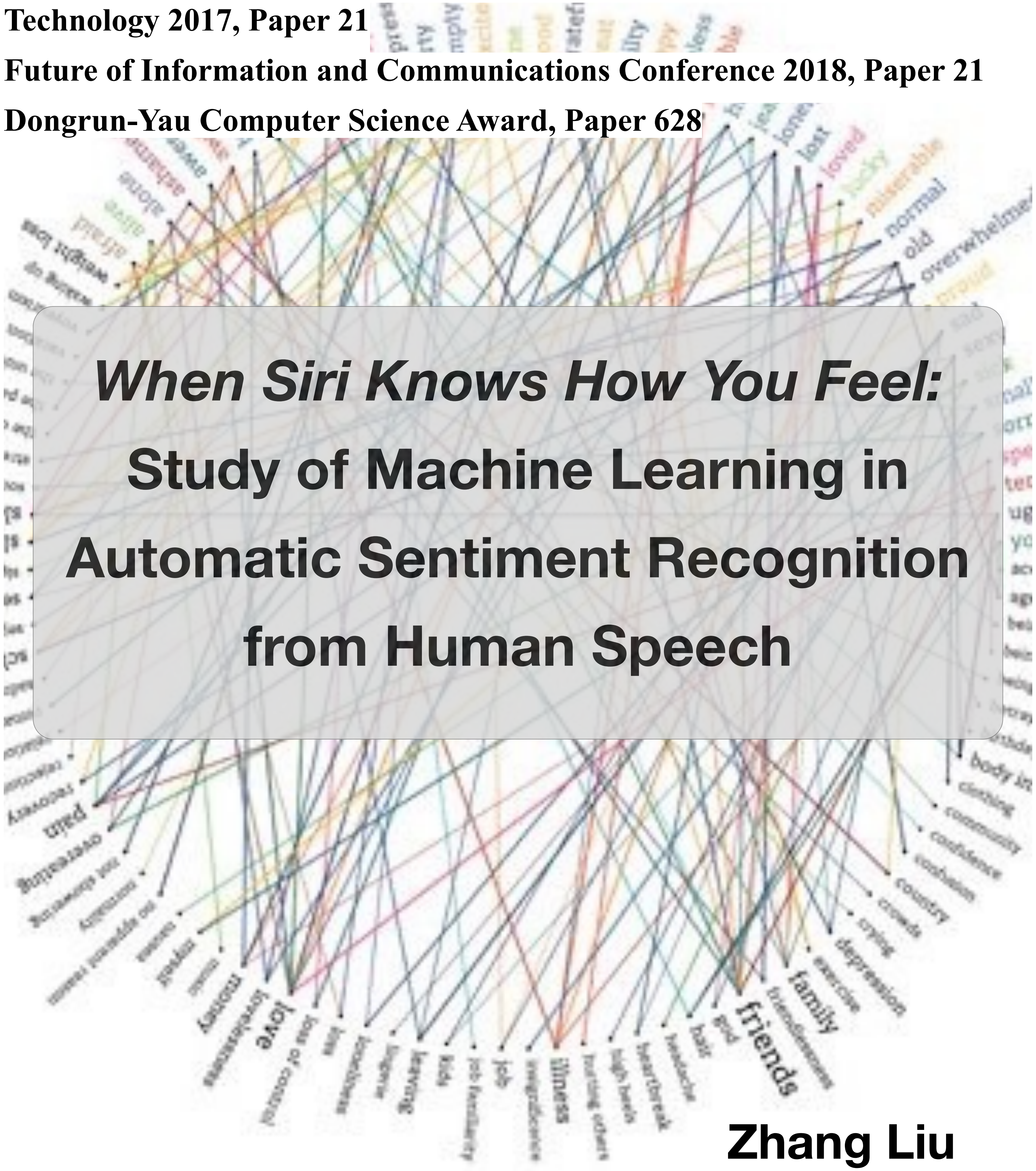
Nanyang Research Program 2016/17, MAE 02

Singapore Science and Engineering Fair 2017, RO005

International Researchers' Club Conference on Science, Engineering, and Technology 2017, Paper 21

Future of Information and Communications Conference 2018, Paper 21

Dongrun-Yau Computer Science Award, Paper 628



When Siri Knows How You Feel: Study of Machine Learning in Automatic Sentiment Recognition from Human Speech

Zhang Liu

Final Product
(as of 10/28/17)

This pdf will include, by order:

- Final Research Paper
- Final Research Poster
- Final Presentation Slides
- My Resume
- Mentor’s Resume
- Research Journal

Future of Information and Communication Conference (FICC) 2018
5-6 April 2018 | Singapore

When Siri Knows How You Feel: Study of Machine Learning in Automatic Sentiment Recognition from Human Speech

Zhang Liu
Anglo-Chinese Junior College
Singapore
ifisay.617@gmail.com

Ng EYK
College of Engineering
Nanyang Technological University (NTU)
Singapore

***Abstract*—Opinions and sentiments are essential to human activities and have a wide variety of applications. As many decision makers turn to social media due to large volume of opinion data available, efficient and accurate sentiment analysis is necessary to extract those data. Hence, text sentiment analysis has recently become a popular field and has attracted many researchers. However, extracting sentiments from audio speech remains a challenge. This project explored the possibility of applying supervised Machine Learning in recognizing sentiments in English utterances on a sentence level. In addition, the project also aimed to examine the effect of combining acoustic and linguistic features on classification accuracy. Six audio tracks were randomly selected to be training data from 40 YouTube videos (monologue) with strong presence of sentiments. Speakers expressed sentiments towards products, films, or political events. These sentiments were manually labelled as negative and positive based on independent judgment of three experimenters. A wide range of acoustic and linguistic features were then analyzed and extracted using sound editing and text mining tools, respectively. A novel approach was proposed, which used a simplified sentiment score to integrate linguistic features and estimate sentiment valence. This approach improved negation analysis and hence increased overall accuracy. Results showed that when both linguistic and acoustic features were used, accuracy of sentiment recognition improved significantly, and that excellent prediction was achieved when the four classifiers were trained, respectively, namely, kNN, SVM, Neural Network, and Naïve Bayes. Possible sources of error and inherent challenges of audio sentiment analysis were discussed to provide potential directions for future research.**

Keywords—Sentiment analysis; natural language processing; machine learning; affective computing; data analytics; speech processing; computational linguistic

I. Introduction

Sentiment analysis is the field of study that analyses opinions, sentiments, appraisals, attitudes, and emotions toward entities and their attributes [1]. Opinions and sentiments are essential to human activities and have a wide variety of applications. As many decision makers turn to social media due to large volume of opinion data available, efficient and accurate sentiment analysis is necessary to extract those data. Business organizations in different sectors use social media to find out consumer opinions to improve their products and services. Political party leaders need to know the current public sentiment to come up with campaign strategies. Government agencies also monitor citizens' opinions on social media. Police agencies, for example, detect criminal intents and cyber threats by analyzing sentiment valence in social media posts. In addition, sentiment information can be used to make predictions, such as in stock market, electoral politics and even box office revenue. Moreover, sentiment analysis that moves towards achieving emotion recognition can potentially enhance psychiatric treatment as emotions of patients are more accurately identified.

Since 2000, researchers have made many successful attempts in text sentiment analysis. In comparison, audio sentiment analysis does not seem to receive as much attention. It is, however, equally significant as text sentiment analysis. Many people in the contemporary society share their opinions using online-based multimedia platforms such as YouTube videos, Instagram stories, TV talk shows and TED talks. It is difficult to manually classify sentiments in them due to the sheer amount of data. With the help of machine automation, we can recognize, with an acceptable accuracy, the general sentiments about certain products, movies, and socio-political events, hence aiding decision-making process of corporations, societal organizations and governments.

This project explored the possibility of using a machine learning approach to recognize sentiments accurately and automatically from natural audio speech in English. In addition, the project also aimed to examine the effect of combining acoustic and linguistic features on classification accuracy. Training data consisted of 150 speech segments extracted from six YouTube videos of different genres. Both acoustic features and linguistic features were examined in order to increase the accuracy of automatic sentiment recognition. Sentiments were categorized into two target classes: positive and negative.

II. Literature Review

There were previous attempts to combine acoustic and linguistic features of speech in sentiment analysis. Chul & Narayanan (2005) [2] explored the detection of domain-specific emotions (negative and non-negative) using language and discourse information in conjunction with acoustic correlates of emotion in speech signals. The database consists of spoken speech obtained from a call center

978-1-5386-2056-4/18/$31.00©2018 IEEE

application. Their results showed that combining all the information, rather than using only acoustic information, improved emotion classification by 40.7% for males and 36.4% for females. This study suggested a comprehensive range of features and provided some insights for my project: acoustic features (Fundamental Frequency (F0), Energy, Duration, Formants), and textual features (emotional salience, discourse information). However, with its speech data collected from a call center, the research focused on emotions in human-machine interactions, rather than in natural human speech.

Another research, Kaushik & Sangwan & Hansen (2013) [3], provided an alternative source of speech data - YouTube videos. In this study, the authors proposed a system for automatic sentiment detection in natural audio streams on social media platform such as YouTube. The proposed technique used Part of Speech (POS) tagging and Maximum Entropy modeling (ME) to design a text-based sentiment detection model. Using decoded Automatic Speech Recognition (ASR) transcripts and the ME sentiment model, the proposed system was able to estimate sentiments in YouTube videos. Their results showed that it was possible to perform sentiment analysis on natural spontaneous speech data despite poor word error rates. This study provided a systematic approach and proved that audio sentiment analysis is possible. It did not, however, include enough acoustic features of audio speech, possibly due to the limitation of document-level analysis.

Ding, *et al*. proposed a holistic lexicon-based approach [4] to solve the problem of insufficient acoustic features by exploiting external evidences and linguistic conventions of natural language expressions. Inspired by above work, a simplified sentiment score model was proposed in this project. The model was useful in sentence level audio speech analysis. The detail of the method will be explained in Section III.

III. METHODOLOGY

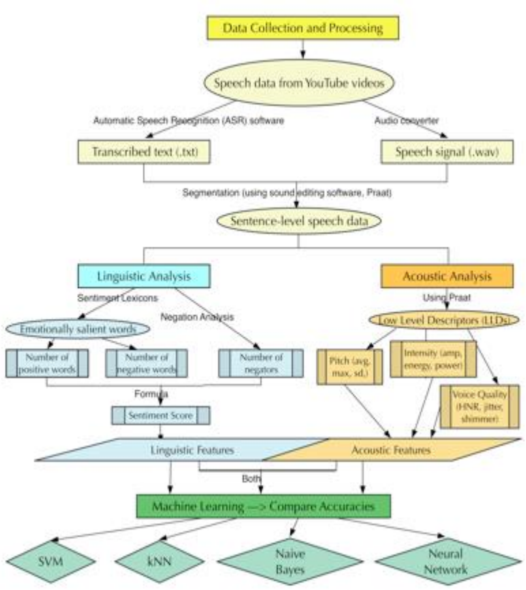


Fig. 1. An overview of the methodology of this project.

A. Database

The speech data used in the experiments were obtained from YouTube, a social media platform. This source was chosen because thousands of YouTube users share their personal opinions or reviews on their channels. Hence, there is a huge amount of accessible speech data containing sentiment valence. More importantly, their ways of speaking are usually closest to natural, spontaneous human speech. Six videos were randomly selected from 40 YouTube videos that had strong presence of negative or positive sentiments. Subject matters included: 1) Product Review; 2) Movie Review; 3) Political Opinion.

During the pre-processing stage, the videos were converted into '.wav' files. Speech transcriptions were generated using the Automatic Speech Recognition (ASR) software, Speechmatics (https://www.speechmatics.com) and checked manually to increase reliability. Each sound file (.wav) was then edited in the vocal toolkit, Praat (http://www.fon.hum.uva.nl/praat/), to match the transcriptions to corresponding sound segments. The TextGrid annotation (as shown in Fig. 2) included 2 tiers, transcription text and numbering, which were useful in keeping track of the data. Meanwhile, the sound file was segmented into smaller sections containing 1 to 5 sentences of relevant meaning and the same sentiment. Each segment was pre-assigned a sentiment label ('negative' or 'positive') based on independent judgment of 3 experimenters so as to minimize bias and subjective errors. There were a total of 150 sound segments (including 70 positive, 80 negative) in the data set. The segmentation process was necessary as most opinion videos contain mixed sentiments.

B. Feature Extraction

1) Linguistic Features

Natural Language Processing toolkit, Orange 3-Text Mining, was used in this stage. Speech transcripts were transformed into lowercase, tokenized into words, and normalized using WordNet Lemmatizer. Part of Speech (POS) tagger was used to label each word as, for instance, a noun, a verb or an adjective, in order to preserve the linguistic function of each word in the sentence. The workflow and text processor parameters were shown in Fig. 3 and 4.

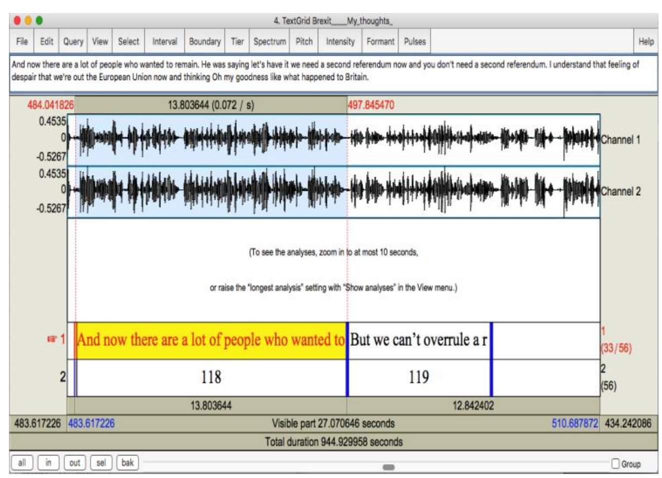


Fig. 2. Using Praat to annotate speech.

978-1-5386-2056-4/18/$31.00©2018 IEEE

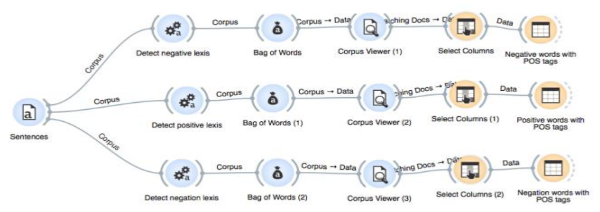
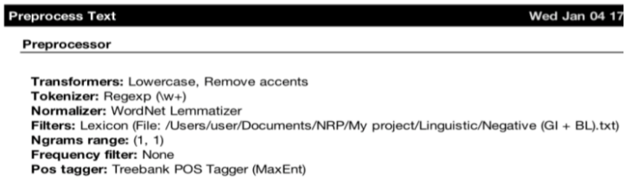


Fig. 3. Orange 3 Text Mining workflow.



Preprocess Text Wed Jan 04 17

Preprocessor

Transformers: Lowercase, Remove accents
Tokenizer: Regexp (\w+)
Normalizer: WordNet Lemmatizer
Filters: Lexicon (File: /Users/user/Documents/NRP/My project/Linguistic/Negative (GI + BL).txt)
Ngrams range: (1, 1)
Frequency filter: None
Pos tagger: Treebank POS Tagger (MaxEnt)

Fig. 4. Text preprocessor parameters.

TABLE I. LIST OF NEGATION CUES

aint	doesnt	havent	lacks	nobody	prevent
arent	dont	havnt	mightnt	none	rarely
barely	doubt	improbable	mustnt	nor	scarcely
cannot	few	isnt	neednt	not	seldom
cant	hadnt	lack	neither	nothing	shant
darent	hardly	lacked	never	nowhere	shouldnt
didnt	hasnt	lacking	no	oughtnt	unlikely
wasnt	werent	without	wouldnt	little	

Textual feature extraction was done by filtering the emotionally salient words (negatively connoted words and positively connoted words). Words in the training corpus were looked up against Harvard General Inquirer and Opinion Lexicon by Bing, Liu [5] to decide if they were negatively or positively connoted. The frequencies of negatively and positively connoted words in each segment were then counted respectively and the numerical values were stored in the training data set.

For negation cues, a similar approach was adopted to look up words against a list of explicit negation cues (compiled manually) as shown in Table I.

Last but not least, a simplified "Sentiment Score" model was proposed to "integrate" all the linguistic features that provide emotion-related information. The Sentiment Score reflected the sentiment of an opinion, with sentiment defined as the valence. For every opinion segment O, a sentiment score, $f(O)$, was calculated using the formula below:

$$f(O) = pos(O) - neg(O) + 2 \times \left(\sum (-1)^{neg\_pos(w)} - \sum (-1)^{neg\_neg(w)}\right),$$

where $pos(O)$ is the number of positive words in the opinion segment, O; $neg(O)$ is the number of negative words in the opinion segment, O; $neg\_pos(w)$ is the number of times each positive word is negated, and similarly, $neg\_neg(w)$ is the sum of the number of times each negative word is negated. Note that $\sum(-1)^{neg\_pos(w)})$ and $\sum(-1)^{neg\_neg(w)}$ were counted manually to give the most reliable values. The following are some advantages of this model:

a) The problem of multiple negation can be solved. When a word is negated twice (with our loss of generality, suppose it is a positive word, as in "can't live without"), the formula will correctly give a positive value that signifies positive sentiment.

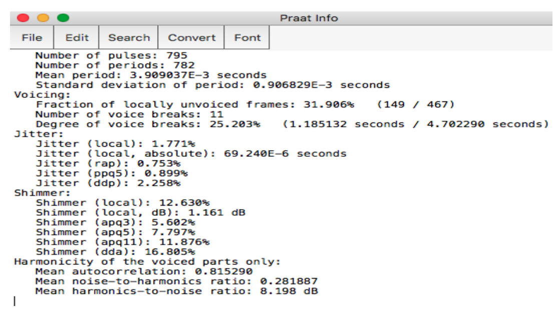
b) It allows semi-automatic negation analysis and has potential to be developed into a fully automated process.

2) *Acoustic Features*

Acoustic features of the sound segments were extracted manually using built-in functions in Praat, as shown in Fig. 5. In order to achieve a more comprehensive representation of the sound, I chose a sufficiently wide range of acoustic features: intensity (amplitude, total energy, mean power), pitch (maximum pitch, average pitch, standard deviation, mean absolute slope), and voice quality (jitter, shimmer, Mean harmonics-to-noise ratio). Considering the inherent differences in pitch between females and males, an attribute "gender" was included to normalize the data.

C. Machine Learning

In the Orange Canvas Software [6], kNN, NN, Naïve Bayes and SVM were used to evaluate the proposed method. Extracted features were sent to appropriate classifier (Fig. 6). Sentiment label (positive, negative) was selected as the target class and the rest of the features as attributes. Stratified 10-folds cross-validation method was used to measure model performance. Hence, each time the dataset was split into 10-folds and one out of ten folds was randomly selected for testing. After multiple experiments, optimal configuration for each classifier was determined and used in the machine learning process.



```
Praat Info
File  Edit  Search  Convert  Font
   Number of pulses: 795
   Number of periods: 782
   Mean period: 3.909037E-3 seconds
   Standard deviation of period: 0.906829E-3 seconds
Voicing:
   Fraction of locally unvoiced frames: 31.906%   (149 / 467)
   Number of voice breaks: 11
   Degree of voice breaks: 25.203%   (1.185132 seconds / 4.702290 seconds)
Jitter:
   Jitter (local): 1.771%
   Jitter (local, absolute): 69.240E-6 seconds
   Jitter (rap): 0.753%
   Jitter (ppq5): 0.899%
   Jitter (ddp): 2.258%
Shimmer:
   Shimmer (local): 12.630%
   Shimmer (local, dB): 1.161 dB
   Shimmer (apq3): 5.602%
   Shimmer (apq5): 7.797%
   Shimmer (apq11): 11.876%
   Shimmer (dda): 16.805%
Harmonicity of the voiced parts only:
   Mean autocorrelation: 0.815290
   Mean noise-to-harmonics ratio: 0.281887
   Mean harmonics-to-noise ratio: 8.198 dB
```

Fig. 5. Extracting acoustic features using Praat.

TABLE II. OPTIMAL CONFIGURATION FOR DIFFERENT CLASSIFIERS

Classifier	Optimal Configurations
k Nearest Neighbors	▪ k = 74 (weighting by distances) ▪ Euclidean (normalize continuous attributes)
Naïve Bayes	▪ Prior: Relative Frequency ▪ Conditional: M-Estimate (parameter = 2.0) ▪ Size of LOESS window = 1.0 ▪ LOESS sample points = 11 ▪ Adjust threshold
Neural Network	▪ Hidden layer neurons = 11 ▪ Regularization factor = 1.0 ▪ Max iterations = 300 ▪ Normalize data
Support Vector Machine	▪ C-SVM (C = 1.00) ▪ Linear Kernel, x- y ▪ Numerical tolerance = 0.0010 ▪ Estimate class probabilities ▪ Normalize data

978-1-5386-2056-4/18/$31.00©2018 IEEE

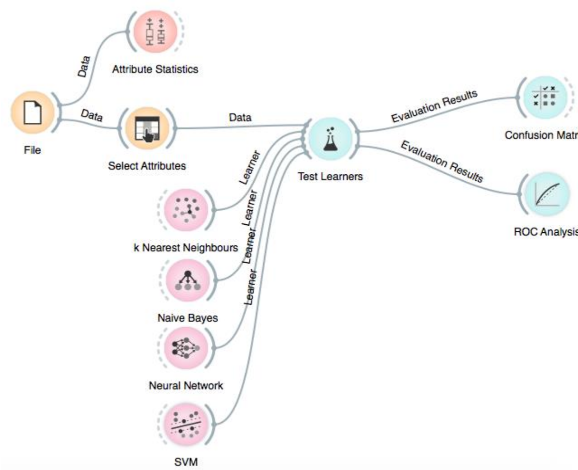


Fig. 6. Illustration for machine learning workflow.

IV. Results and Discussions

A. Results Analysis

The evaluation will be focused on Area Under the ROC Curve (AUC) as it has "better statistical foundations than most other measures" [7], ROC Area Benchmark: 1.0: perfect prediction; 0.9: excellent prediction; 0.8: good prediction; 0.7: mediocre prediction; 0.6: poor prediction; 0.5: random prediction; <0.5: something wrong. As shown in Table 2, accuracy improved significantly when both acoustic and linguistic features were used, instead of only acoustic features or only linguistic features. When both acoustic and linguistic features were extracted, excellent classification of sentiments was achieved when the four classifiers were trained, with kNN, SVM and Neural Network having higher accuracies. The shapes of ROC curves for these four classifiers resembled the shape of ROC curve for excellent prediction (Fig. 7 and 8) [8].

TABLE III. AUC when Different Classifiers & Features are used

Classifier	**AUC (acoustic features only)**	**AUC (linguistic features only)**	**AUC (Both acoustic features & linguistic features)**
kNN	0.8750	0.8420	0.9321 > 0.9
Naïve Bayes	0.7964	0.8348	0.8929 ≈ 0.9
Neural Network	0.9018	0.8384	0.9304 > 0.9
SVM	0.8589	0.8607	0.9429 > 0.9

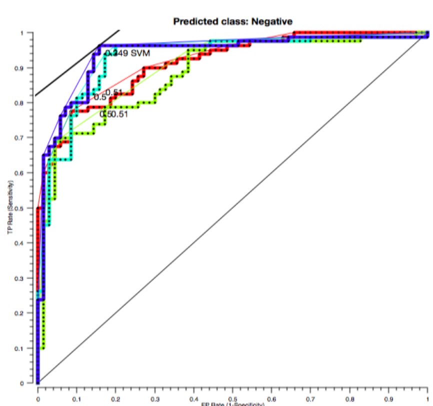


Fig. 7. ROC curves for different classifiers when both acoustic and linguistc features were used.

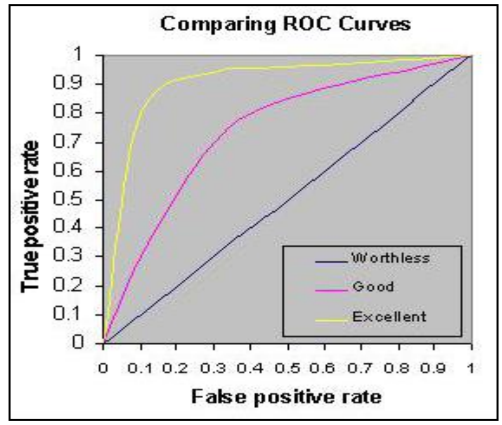


Fig. 8. Shapes of ROC curves indicate different levels of accuracy.

B. Limitations and Sources of Error

The speech corpus might not be large enough. There has to be sufficient representation of the different sentiments and different speech types in order for more comprehensive learning, and hence more accurate recognition. A standard sentiment analysis database could be used as a benchmark to compare the accuracy of this linguistic/acoustic model against other models. The average Word Error Rate of the Automated Speech Recognition (ASR) software Speechmatics is reported to be 33%. Hence, transcription errors might still be present even after manual check, affecting the accuracy of linguistic analysis. Negation analysis is subject to human error and there might be inaccurate detection of context-specific meanings of polysemes[1].

C. Inherent Challenges

Sentiment analysis is a challenging task due to ambiguities in language, such as subtlety, concession, manipulation, sarcasm and ironies in speech. To address this problem, an accurate conclusion might entail examination of other features such as physiological symptoms (blood pressure, etc.) and facial expressions. Although inaccuracies arising from ambiguities could be minimized by analyzing data from multiple dimensions, cultural differences and multilinguality[2] further complicate the process. Due to differences in cultural backgrounds, the ways people express their sentiments vary among individuals. (For example, the way a Japanese expresses a sentiment differs from the way an American expresses the same sentiment). Moreover, sentiment expressions depend on contexts of speech, and hence vary even for the same person at different times. In addition, the speakers might constantly change subject or compare with another subject, which might be hard to detect.

There are also issues with mutual interpretability. Interjections that express feelings (such as "urggghhh") might be deemed as irrelevant by the machine. It might be hard, if not impossible, for the machine to "master" contextual knowledge such as some exophoric references[3] to historical

[1] A word or lexical unit that has several or multiple meanings

[2] Multilinguality is a characteristic of tasks that involve the use of more than one natural language. (Kay, n.d.)

[3] *Exophoric reference is* referring to a situation or entities outside the text. (University of Pennsylvania, 2006)

978-1-5386-2056-4/18/$31.00©2018 IEEE

figures ("the German dictator", which refers to Hitler). The issue becomes more significant when dialects are used. For example, the negation analysis is based on Standard English usage, which might not be useful for other varieties of English. Speakers of certain dialects like African American Vernacular English (AAVE) usually employ double negatives to emphasize the negative meaning.

V. CONCLUSION AND FUTURE WORK

In this study, we have built a machine learning model combining acoustic and linguistic features. As the results have shown, this model has significantly higher accuracy than models with only acoustic or only linguistic features. Under this model, excellent prediction can be achieved. Although limitations and challenges are real and a considerable amount of manual work is necessary, the positive results of this study have clearly suggested to the possibility of achieving a fully automated audio sentiment analysis in future.

Based on the limitations and challenges discussed in Section IV, the following three main directions of research are proposed:

1) From audio sentiment analysis towards video sentiment analysis by incorporating facial expression features, and further towards multi-dimensional sentiment analysis by incorporating physiological features such as blood pressure and heart rate.

2) From *semi*-automatic sentiment analysis towards *fully* automatic sentiment analysis, by reducing the amount of manual processing of data.

3) From *sentiment* recognition towards *emotion* recognition, by enabling classification of specific emotions such as fear, anger, happiness, sadness.

Artificial Intelligence (AI) is becoming an increasingly interdisciplinary field. To achieve the above research goals, cross-discipline cooperation is crucial. Solutions to the challenges of language/emotion recognition and understanding can be inspired by diverse fields from mathematics and sciences, which provide us with quantitative methods and computational models, to humanities and fine arts, which shed light on qualitative analysis and feature selection. From a neuroscience perspective, learning about how the human brain perceives and processes sentiments and emotions might inspire a better machine learning architecture for sentiment prediction. Mathematical modelling could be useful as well: the high complexity of emotions should be captured more comprehensively by mapping the emotion of each utterance in multi-dimensional vector space. Linguistics theories also imply that language is meaningless without context (the socio-cultural background of the speaker, the conversation setting, and the general mood). It is a timely reminder for Natural Language Processing (NLP) researchers to go beyond *content analysis* – dissecting language as an isolated entity only made up of different parts of speech – and aim for "*context analysis*". Without being context-aware, AI will only be machines with "high Intelligence Quotient (IQ)" but "low emotional intelligence quotient (EQ)". Emotion theory in drama and acting also provides some insights for developing affective, sentient AI. For example, emotions can be conveyed through subtle means such as silence, cadence, and paralinguistic features (kinesics, i.e. body language and proxemics i.e. use of space etc.). This will give us directions in selecting and extracting features salient to sentiment.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my project supervisor Professor Eddie Ng for being open-minded about my project topic, without which I could not have been able to delve deep into my field of interest. His insightful suggestions and unwavering support has guided me through doubts and difficulties.

REFERENCES

[1] Liu, Bing. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press.

[2] Chul Min Lee & Shrikanth S. Narayanan. (2005). Toward Detecting Emotions in Spoken Dialogs. *IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 13, NO. 2, MARCH 2005*.

[3] Kaushik, Lakshmish & Sangwan, Abhijeet & Hansen, John H.L. (2013). A Holistic Lexicon-Based Approach to Opinion Mining. *IEEE*.

[4] Ding, Xiaowen, Bing Liu, and Philip S. Yu. A Holistic Lexicon-Based Approach to Opinion Mining. In *Proceedings of the Conference on Web Search and Web Data Mining (WSDM- 2008)*. 2008.

[5] Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA

[6] Demšar, J., Curk, T., & Erjavec, A. (2013). Orange: Data mining toolbox in Python. *Journal of Machine Learning Research*, *14*, 2349−2353.

[7] Unknown. Cornell University. (2003). Retrieved from: https://www.cs.cornell.edu/courses/cs578/2003fa/performance_measures.pdf

[8] Tape, Thomas G. (n.d.). *Interpreting Diagnostic Tests*. University of Nebraska Medical Center. Retrieved from: http://gim.unmc.edu/dxtests/roc3.htm

978-1-5386-2056-4/18/$31.00©2018 IEEE

When Your Siri Knows How You Feel:
Application of Machine Learning in Automatic Sentiment Recognition from Human Speech

Zhang Liu

Introduction

Aim

- Due to the **rise of social media**, people express opinions about almost every aspect of their lives online.
- This project explores the possibility of using supervised Machine Learning to recognize sentiments ***accurately*** and ***automatically*** from natural audio speech in English.

Significance

Business organizations improve their products and services based on consumer opinions.

Politicians need to know prevailing public opinions so as to design more effective campaign strategies.

Psychiatric treatment can be enhanced as emotions of patients are accurately identified.

All of the above, however, involve **a huge amount of data**. Hence, there is a need for automatic Sentiment Analysis performed by machine.

Methodology

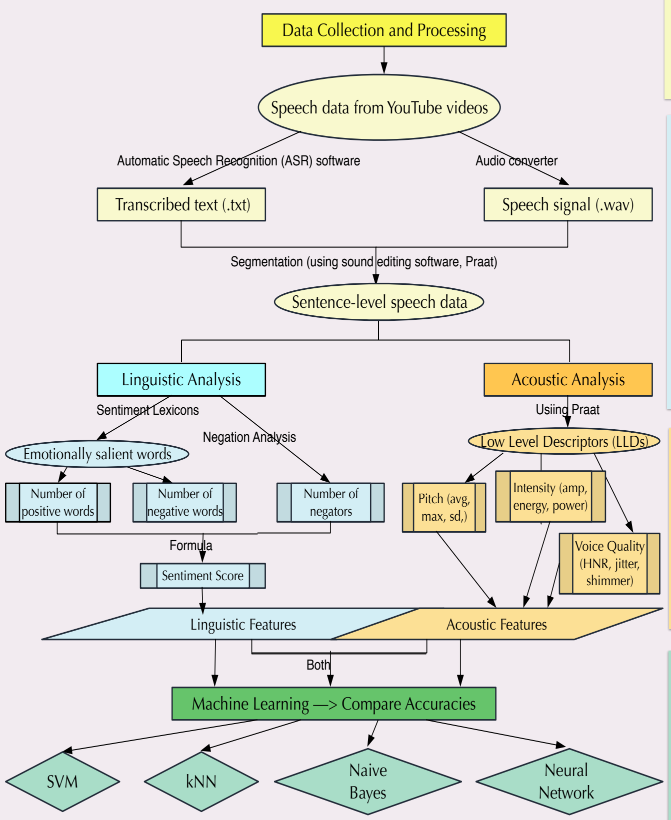


Figure 1 *Overview of Methodology*

Data Collection

- **166** audio segments from 6 YouTube videos
- **3** genres: Product Review, Movie Review, and Political Opinion
- Classification/Label: positive and negative

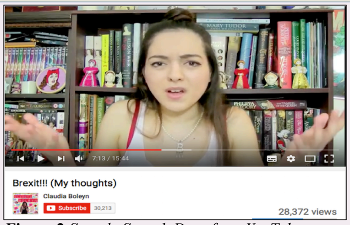


Figure 2 *Sample Speech Data from YouTube.com*

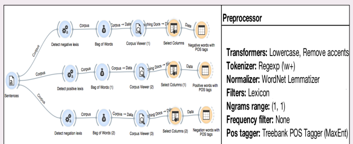
Linguistic Analysis

For each opinion, *O*, find:

- **Pos(*O*)**, the number of word connoting positive sentiment
- **Neg(*O*),** the number of word connoting negative sentiment

For each emotionally salient word, *w*, find:

- **Neg_Pos(*w*)** and **Neg_Neg(*w*)**, the number of negation(s)
- Hence, calculate the ***'Sentiment Score'***.



$$f(O) = pos(O) - neg(O) + 2 \times \left(\sum(-1)^{neg\_pos(w)}\right) - \sum(-1)^{neg\_neg(w)}\right)$$

Figure 3 *Linguistic Analysis: 'Sentiment Score'*

Acoustic Analysis

- **Voice Quality:** jitter, shimmer, mean Harmonics-to-Noise Ratio (HNR)
- **Pitch:** mean absolute slope, average pitch, maximum pitch, standard deviation
- **Intensity:** amplitude, total energy, mean power

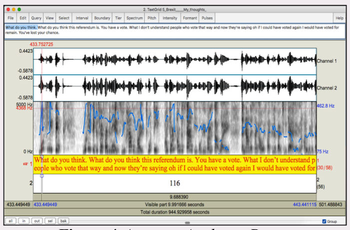


Figure 4 *Acoustic Analysis: Praat*

Machine Learning

- 10-fold cross-validation
- **Compare and evaluate** the accuracy of sentiment classification, when k Nearest Neighbor (kNN), Naive Bayes, Neural Network, and Support Vector Machine (SVM) are used as the classifier respectively.

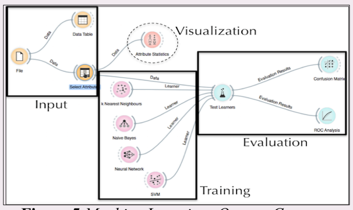


Figure 5 *Machine Learning: Orange Canvas*

Results & Discussion

- ***AUC*** → "better statistical foundations than most other measures"
- When **both acoustic and linguistic features** are extracted, accuracy of sentiment classification has significantly improved.
- **Excellent prediction (AUC > 0.9)** is achieved, with **kNN** and **Neural Network** having a better performance. The shapes of the ROC curves also indicate excellent prediction.

Classifier	AUC (acoustic features only)	AUC (linguistic features only)	AUC (Both acoustic features & linguistic features)
kNN	0.8750	0.8420	0.9321 > 0.9
Naïve Base	0.7964	0.8348	0.8929 ≈ 0.9
Neural Network	0.9018	0.8384	0.9304 > 0.9
SVM	0.8589	0.8607	0.9232 > 0.9

Figure 6 *AUC when different classifiers & features are used*

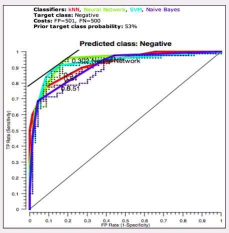


Figure 7 *ROC curves generated using Orange Canvas*

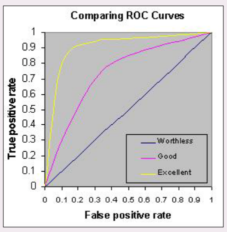


Figure 8 *ROC Analysis Benchmark*

Potential Sources of Error

- Speech corpus is not large enough. Hence, it might result in inaccuracy due to over-fitting.
- Error might occur when generating the transcriptions using Automatic Speech Recognition (ASR) software, even after human check.
- Negation analysis is subjected to human errors.

Inherent Challenges

- Language can be ambiguous: subtlety, concession, manipulation, sarcasm, and irony.
- Language is culture and context specific: cultural differences and individual variations.
- Speakers might change the subject or make comparisons with another subject.
- Machine might not have sufficient historical/cultural knowledge to understand an opinion expressed by the speaker.

Conclusion

As the results have show, with **both acoustic and linguistic features** as input data, excellent prediction can be achieved if **Neural Network** is used to classify the sentiments. Although limitations and challenges are real and a considerable amount of manual work is necessary, the positive results of this study have clearly pointed to the **possibility of achieving fully automated audio sentiment analysis** in future.

Future Research

3 main directions of research:

- ***Audio*** Sentiment Analysis (+ Facial expression features) → ***video*** Sentiment Analysis (+ physiological features) → ***multi-dimensional*** Sentiment Analysis
- ***Semi***-automatic → ***fully*** automatic
- ***Sentiment*** recognition → ***emotion*** recognition

ALL figures are self-drawn illustrations or workflow screenshots, EXCEPT for **Figure 8**.

REFERENCES: *Figure 8.* Tape, Thomas G. (n.d.). *Interpreting Diagnostic Tests*. University of Nebraska Medical Center. Retrieved from: http://gim.unmc.edu/dxtests/roc3.htm

When Your Siri Knows How You Feel :

Application of Machine Learning in Automatic Sentiment Analysis in Human Speech


her

Significance

- Business organizations
- Politicians
- Psychiatric treatment

All of the above, however, involve **a huge amount of data.**

Hence, there is a need for automatic Sentiment Analysis performed by machine.

Aim

- To use supervised Machine Learning to recognize sentiments ***accurately*** and ***automatically*** from natural audio speech in English
- To investigate if a combination of acoustic and linguistic features improves accuracy

Can you identify the speaker's sentiment on Brexit?

Methodology

Sentiment Score

$$f(O) = pos(O) - neg(O) + 2 \times \left(\sum (-1)^{neg\_pos(w)}\right) - \sum (\text{-}1)^{neg\_neg(w)}),$$

- *For every opinion, O, f(O) is its sentiment score.*
- *pos(O), number of positive words*
- *neg(O), number of negative words*
- *neg_pos(w), number of negations to a positive word*
- *neg_neg(w), number of negations to a negative word*

I don't think it's not scary, and I'm kind of worried.

Classifier	Optimal Configurations
k Nearest Neighbors	▪ k = 74 (weighting by distances) ▪ Euclidean (normalize continuous attributes)
Naïve Bayes	▪ Prior: Relative Frequency ▪ Conditional: M-Estimate (parameter = 2.0) ▪ Size of LOESS window = 1.0 ▪ LOESS sample points = 11 ▪ Adjust threshold
Neural Network	▪ Hidden layer neurons = 11 ▪ Regularization factor = 1.0 ▪ Max iterations = 300 ▪ Normalize data
Support Vector Machine	▪ C-SVM (C = 1.00) ▪ Linear Kernel, x- y ▪ Numerical tolerance = 0.0010 ▪ Estimate class probabilities

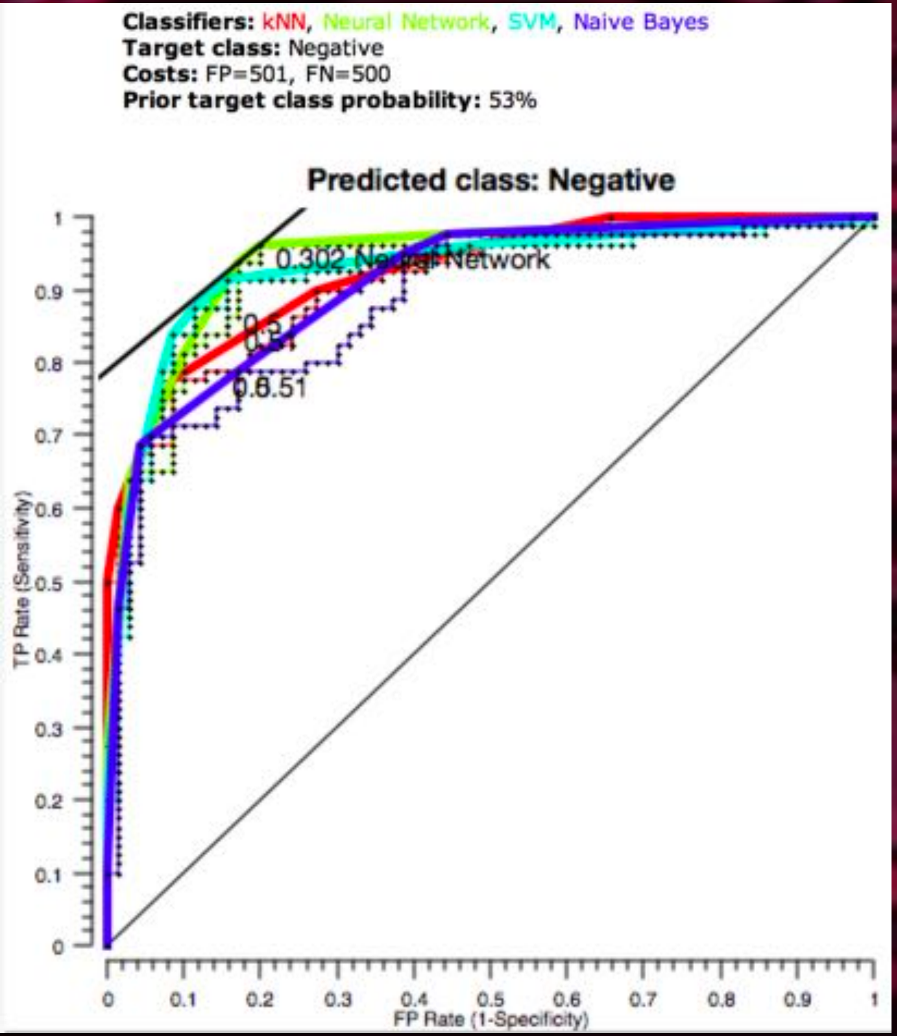
Results & Conclusion

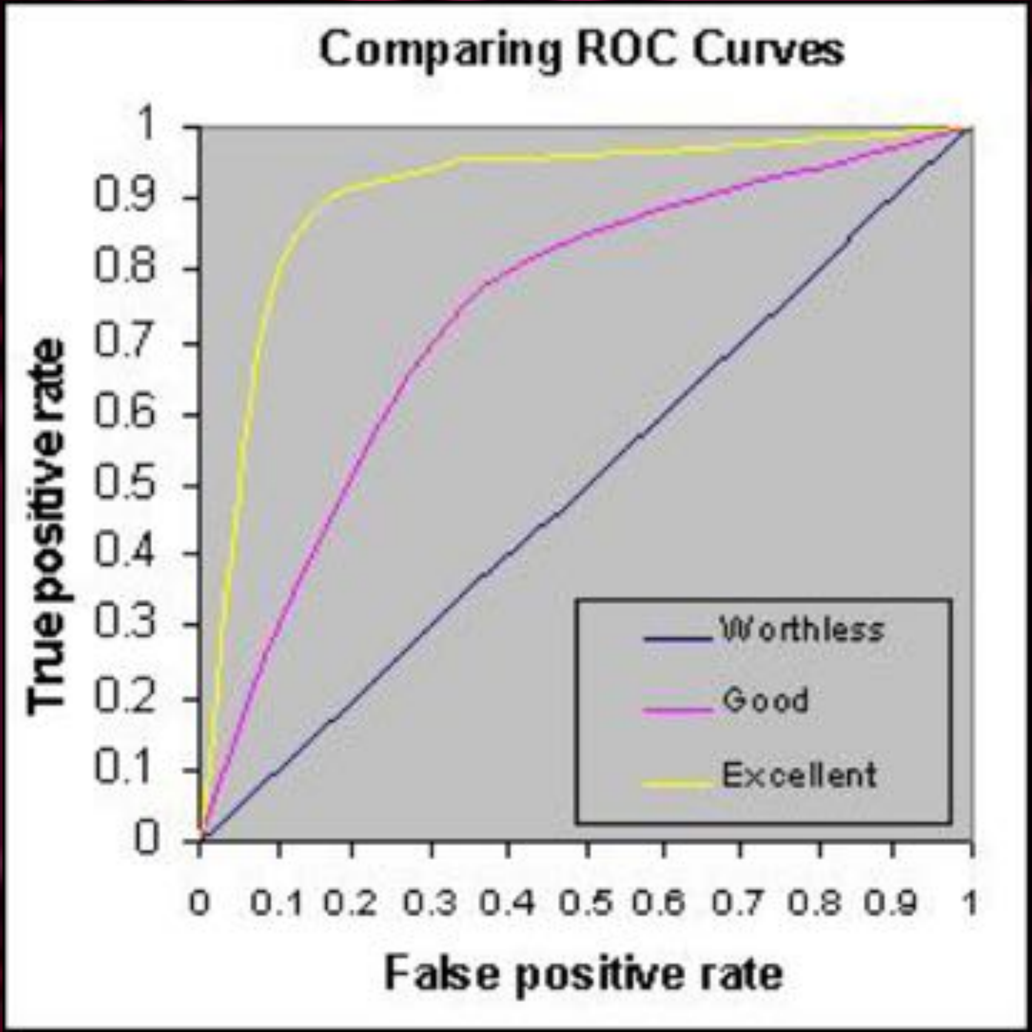
- *Area Under the Receiver Operating Characteristic (ROC) Curve (AUC):* “better statistical foundations than most other measures”

- Accuracy has significantly improved when both acoustic and linguistic features are extracted.
- Excellent prediction (AUC > 0.9)

Classifier	AUC (acoustic features only)	AUC (linguistic features only)	AUC (Both acoustic features & linguistic features)
kNN	0.8750	0.8420	0.9321 > 0.9
Naïve Base	0.7964	0.8348	0.8929 ≈ 0.9
Neural Network	0.9018	0.8384	0.9304 > 0.9
SVM	0.8589	0.8607	0.9232 > 0.9

- ROC curves indicate excellent prediction.





Discussion

- Language can be **ambiguous**: subtlety, concession, manipulation, sarcasm, and irony.
- Language is **culture and context specific.**
- Speakers might change the subject or make comparisons with another subject.

3 Directions of Future Research

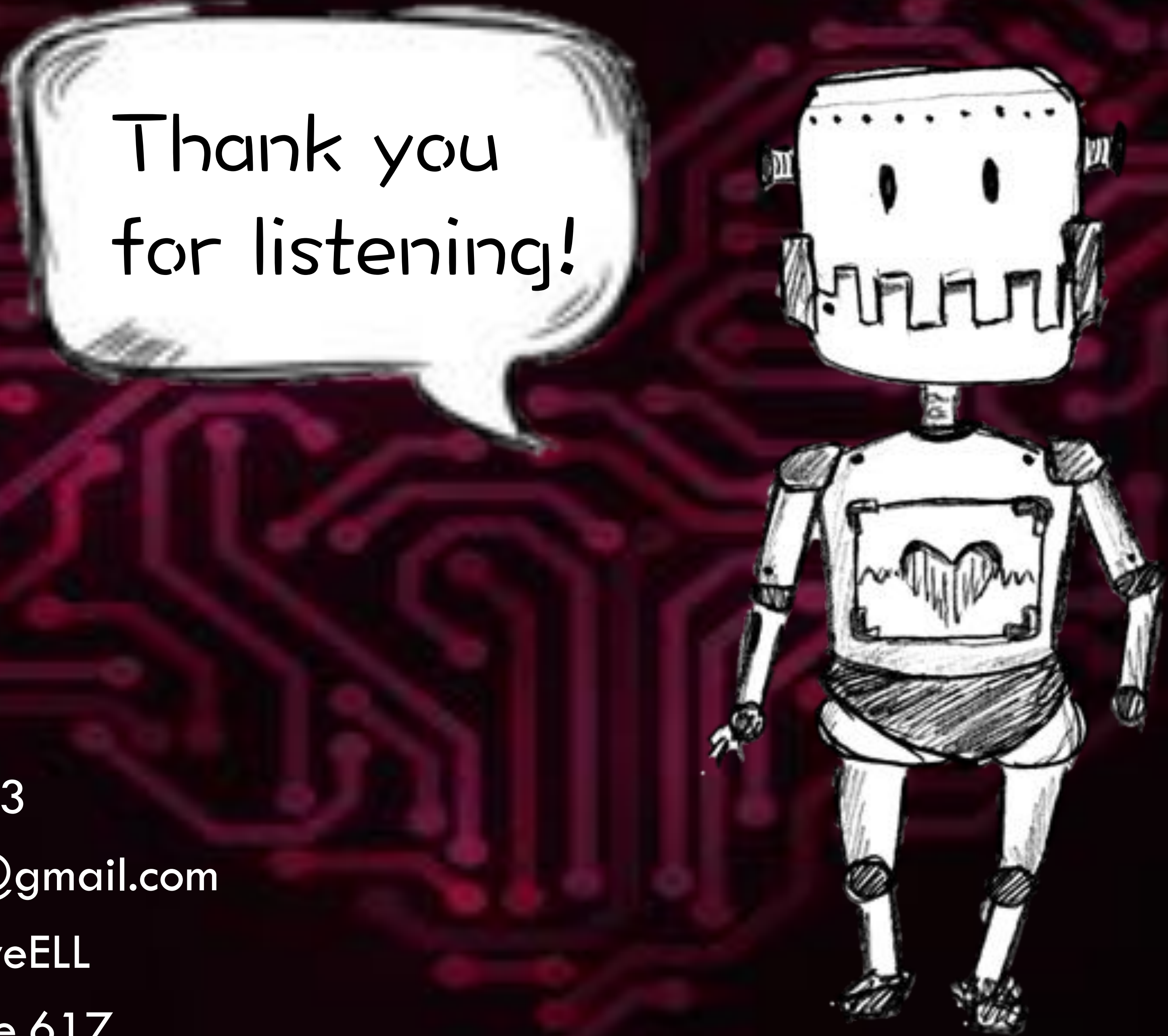
Audio Sentiment Analysis

Semi-automatic

Sentiment Recognition

Acknowledgement

I would like to express my sincere gratitude to my project supervisor, Professor Ng Yin Kwee (NTU), for his insights and open-mindedness. I am also grateful for the support from my teachers, family and friends



Zhang Liu (ACJC)

Tel: +65 8515 6883

Email: zhangliu.sci@gmail.com

Twitter: @ireallyloveELL

Instagram: @corrine.617

Resume

ZHANG LIU

119 Dover Road, ACSI Boarding School, Singapore, S139656

(65) 85156883

zhangliu.sci@gmail.com

Education

Anglo-Chinese Junior College, Singapore 2016-2017

Swiss Cottage Secondary School, **Singapore** 2014-2015

No. 13 Middle School, **Urumqi, Xinjiang, China** 2010-2013

Activities

- **Organizer**, Swiss Cottage Secondary School Secondary One Orientation (2014)
- **Organizer**, Student Council Recruitment (2014)
- **Participant**, Science and Technology Endowment Program (STEP) - National University of Singapore (NUS) Sunburst Brain Camp (2017)
- **Participant**, National University of Singapore (NUS) International Olympiad in Informatics (IOI) Workshop (2014)

Leadership

- **Vice-President (Talent Development)**, Science and Mathematics Council (SMC), Anglo-Chinese Junior College (2016-2017)
- **President**, Mathematics and Computing Club, Anglo-Chinese Junior College (2016-2017)
- **Overall In-Charge (IC)**, Open House CCA Showcase, CCA Exhibition & CCA Recruitment (2017)
- **Event IC**, Vietnam Educational Program – SOS Village (2016)
- **Assistant Event IC**, Heyday (2016)
- **Assistant Event IC** and **Computer Science Question Setter**, International C. B. Paul Quiz (2017)
- **Committee Chairman**, International Biomedical Quiz (2016 & 2017), International C. B. Paul Quiz (2016 & 2017), International Young Whizzes Challenge (2016)
- **Prefect**, Student Council Training and Development Committee, Swiss Cottage Secondary School (2014-2015)
- **Team Leader**, RoboCup (2015)**,** SMU Youth Innovation Challenge (2015), National Junior Robotics Competition (2014)

Research Experience

1. **Research Student** **April 2016 - April 2017**
   - College of Engineering (Mechanical Aerospace Engineering), Nanyang Technological University, Singapore Mentor: Associate Professor Ng, Yin Kwee, Eddie (http://www.mae.ntu.edu.sg/aboutus/FacultyandStaff/Faculty/Pages/mykng.aspx)
   - Research topic: Machine Learning and Sentiment Analysis

2. **Research Student** **Feb 2017- April 2017**
    - Yale-NUS College, Singapore Mentor: Assistant Professor Jean Liu (https://www.yale-nus.edu.sg/about/faculty/jean-liu/)
    - Research topic: the neuroscience and psychology of love
3. **Research Intern** **June 9th to 20th, 2014**
    - Centre for Research and Applied Learning in Science, Science Centre Singapore Research topic: build an automated lighting system with Arduino and sensors

Projects and Publications

1. **"When Siri Knows How You Feel: Application of Machine Learning in Automatic Sentiment Recognition in Human Speech"**
    - Presented in International Researchers' Club Conference on Science Engineering and Technology (2017) and won the Best Paper Award
    - Paper accepted and will be presented at the SAI Future of Information and Communications Conference (IEEE technically sponsored) (2018)
    - Presented at Nanyang Research Program (2016/17)
    - Presented at Singapore Science and Engineering Fair (2017)
2. **Effects of Broad-Spectrum Antibiotics on The Clinical Outcomes of Pneumonia** (2017)
    - Presented at Beijing Youth Science Creation Competition (2017)
3. **Multi-Perspective Research on Autism Spectrum Disorder** (2017)
    - Presented at the Science and Technology Endowment Program (STEP) - National University of Singapore (NUS) Sunburst Brain Camp (2017)
4. **Neuroscience of Love**
    - Published in the *Brain Book,* Chapter 6 (2017)
5. **CRADLE Automatic LED Lighting System** (2014)

Honors and Awards

- **Best Paper Award**, International Researchers' Club Conference on Science, Engineering and Technology (2017)
- **Gold**, Nanyang Research Program (2016/17)
- **Silver,** Beijing Youth Science Creation Competition (2017): **Represented Singapore** at international event
- **Bronze**, Singapore Science and Engineering Fair (2017) (highest award for Robotics)
- **Bronze**, Singapore National Olympiad in Informatics (2017 & 2016 & 2015)
- **Silver**, Singapore Mathematics Olympiad (Senior) (2014)
- **Honourable Mention**, Singapore Mathematics Olympiad (Open) (2016)
- **Distinction**, Australian Mathematics Competition (2016)
- **Distinction**, Computational and Algorithmic Thinking (CAT) (2016)
- **Top 20 Runners**, Anglo-Chinese Junior College Intra-School Cross Country (2016)
- **Represented the school** in National Inter-School Cross Country Championship (2015)

- **1st Runner-up**, Swiss Cottage Intra-School Cross Country 'B' Division Girls' Championship (2015 & 2014)
- **High Distinction**, International Competitions and Assessments for Schools (ICAS) in Mathematics (2014)
- **Distinction**, International Competitions and Assessments for Schools (ICAS) in Science (2014)
- **Academic Book Prize for Top in Standards** (2014)
- **Academic Book Prize for Top in Additional Mathematics** (2014)
- **Top Student**, 2nd Preliminary Examinations (2013)
- **Third Prize**, China National Olympiad in Informatics in Province (NOIP) (2012)

Membership

- **Member**, International Researchers' Club (IRC) Singapore (2017-)
- **Member**, ManiAC student journalist group, Anglo-Chinese Junior College (2016-2017)
- **Member**, Science and Technology Club, Swiss Cottage Secondary School (2014-2015)

Areas of Expertise and Interest

1. ***Programming languages:*** Proficient: C/C++; Beginner: Python and JavaScript
2. ***Software / tools:*** Proficient: Statistical Package in Social Sciences (SPSS), Praat, Orange Canvas, Arduino, RoboLab, Mindstorm; Beginner: MATLAB
3. ***Language skills:*** Written and Verbal: English and Mandarin Chinese
4. ***Research interests:*** Computer Science, Natural Language Processing, Sentiment Analysis, Artificial Intelligence, Machine Learning, Affective Computing, Cognitive Science, Neuroscience, Computational Linguistics, Psycholinguistics, Sociolinguistics, Chemical Linguistics
5. ***Personal interests:*** language games and puzzles, long distance running, hiking, sci-fi movies and novels, poetry and literature, psychological thrillers, detective novels, meta-physics theories, Xiao (Chinese Flute), Indie music, MBTI analyses

Name: NG YIN KWEE, Eddie (PI)

Employment History

Name of Organization	Position Held	Period of Employment
Exxon Tankers Fleet	Marine Engineer	1983-1985
MPE-NTU	Lecturer	1992-1996
MPE-NTU	Senior Lecturer	1997-1998
MPE/MAE-NTU	Associate Professor	1999 -
MAE-NTU	Assistant Chair (Alumni)	Mar. 2008-31 Dec. 2012
National University Hospital (NUH)	Adjunct NUH Scientist	Sep. 2006 -

Academic Qualifications

Qualifications obtained & Class of Honors	Name of University	Year of Award
BEng, 1st CL	Uni. of Newcastle upon Tyne	1988
PhD	Uni. of Cambridge	1992
Postgraduate Diploma in Teaching Higher Education	NIE-NTU	1995

Professional Qualifications

Fellow, Cambridge Commonwealth Society since 2003 by Cambridge Commonwealth Trust

Professional Services

No	Engaged Task	Occasion/Event	When (Period)
1	**Editor-in-Chief (Lead)**	Journal of Mechanics in Medicine and Biology (SCI-ISI index)	Since 2007
2	**Editor-in-Chief (Founding)**	Journal of Medical Imaging and Health Informatics (SCI-ISI index)	Since 2011
3	**Strategy Associate Editor-in-Chief**	World Journal of Clinical Oncology	Since 2010
4	Regional Editor	Computational Fluid Dynamics Journal	Since 1998
5	Associate Editor	Int. Journal of Rotating Machinery	Since 2004
6	Associate Editor	Chinese Medical Journals, Chinese Journal of Medicine	Since 2005
7	Associate Editor	Open Medical Informatics Journal	Since 2007
8	Associate Editor	Open Numerical Methods Journal	Since 2008
9	Associate Editor	Journal of Healthcare Engineering	Since 2009
10	Associate Editor	International Journal of Breast Cancer	Since 2010
11	Editor (http://www.aspbs.com/jcp/)	Journal of Scientific Conference Proceedings, ASP	Since 2011
12	Asian Editor (http://www.aspbs.com/jbns/)	Journal of Bionanoscience, ASP	Since 2011
13	Invited lecture speaker	9th Int. Quantitative Infrared Thermography Conf., (Poland)	2-5 July 2008
14	Keynote lecture # K03	16th International Conf. on Mechanics in Medicine & Biology, (Pittsburgh, USA)	23-25, July 2008
15	Keynote lecture # K06	2009 Int. Sym. on Biomechanics cum Annual Conference of Biomedical Engg, and workshop in NCKU, Taiwan	10-14, Dec. 2009
16	Plenary lecture # P18	2010 Int. Conf. on Medical Physiology (Cambridge Univ.)	23-25, Feb. 2010
17	Guest Editor	Int. Journal of Medical Systems (theme: D2H2)	May 2010
18	Guest Editor	Int. Journal of Computer Applications in Technology	2000 & 2004 each
19	Invited lecture speaker	2nd International Workshop on Biomedical Engg. & Biomechanics, 2011 (Xian, China)	12-14 Aug. 2011
20	Invited keynote speech speaker with Honorarium and Accommodation of USD3k	IEEE Conference on Biomedical Engineering and Biotechnology, held in Macau, China.	May 28-30, 2012
21	Invited lecture speaker	Health Engineering 2012, UTM, Malaysia	Nov 29-Dec 1, 12
22	Invited Opening keynote speech speaker with Honorarium and	National Conference of Biomechanical Engineering, Vitoria, Brazil	April 25-28, 2013

	Accommodation		
23	Invited Opening keynote speech speaker with Honorarium and Accommodation	2nd Conference on Biomedical Engineering and Biotechnology, Wuhan, China http://www.icbeb.org/photo.aspx?picture=1.	Oct. 11-13, 2013
24	Keynote lecture (invited)	IEEE 5th Image Processing, Image Analysis & Real-time Imaging Symposium cum 2nd Symposium on Acoustic, Speech & Signal Processing, UniMAP, Perlis, Malaysia	30 April 2014
25	Invited Opening keynote speech speaker	3rd Conference on Biomedical Engineering and Biotechnology, Beijing, China	Sept. 25-28, 2014
26	Invited Opening keynote speech speaker	4th Conference on Biomedical Engineering and Biotechnology, Shanghai, China	18 – 20 Aug 2015
27	Invited Opening keynote speech speaker	2nd International Conference on Computational Methods in Engineering and Health Sciences, Malaysia.	19 – 20 Dec 2015
28	Invited Opening keynote speech speaker with Accommodation	Quantitative InfraRed Thermography Conference, Gdansk, Poland	04 – 08 Jul 2016
29	Invited Opening keynote speech speaker	5th Conference on Biomedical Engineering and Biotechnology, Hangzhou, China	01 – 04 Aug 2016
30	Invited Opening keynote speech speaker	3rd International Conference on Computational Methods in Engineering and Health Sciences, Kitakyushu city, Fukuoka, Japan	17 – 18 Dec. 2016
31	Invited Opening keynote speech speaker	6th Conference on Biomedical Engineering and Biotechnology, Guangzhou, China	17 – 20 Oct 2017

Selected 10 publications: to-date, Web of Science SCI-IF international journals (299), int. conf. proceedings (100), books (14) & textbook chapters (100). Total citation > 3300; H-Index: 35

Ng, E. Y-K. and Sudharsan, N.M. "An Improved 3-D Direct Numerical Modelling and Thermal Analysis of a Female Breast with Tumour", **International Journal of Engineering in Medicine**, Vol. 215, No. 1, **(2001),** Pp. 25 -- 37.
Ng, E. Y-K. and Sudharsan, N.M. "EFFECT OF BLOOD FLOW, TUMOUR AND COLD STRESS IN A FEMALE BREAST: A Novel Time-accurate Computer Simulation", **International Journal of Engineering in Medicine**, Vol. 215, No. H4, **(2001),** Pp.393-404.
Ng, E. Y-K., Kaw, G.J.L, and Chang, W.M., "Analysis of IR Thermal Imager for Mass Blind Fever Screening", **Microvascular Research,** Vol. 68, No: 2, Reed Elsevier Science, Academic Press, USA. **(2004),** Pp. 104 -- 109.
Ng, E. Y-K. and Sudharsan, N.M., "Numerical Modelling in Conjunction with Thermography as an Adjunct Tool for Breast Tumour Detection", **BMC Cancer,** 4(17), Medline Journal **(2004),** 1--26. IF ≈ *2.9.*
Ng, E. Y-K, "Is Thermal Scanner Losing its Bite in Mass Screening of Fever due to SARS?", **Medical Physics,** American Association of Physicists in Medicine **(2005),** 32(1), 93--97. *IF ≈ 2.5.*
Ong, M.L. and **Ng**, E. Y-K, "A Global Bioheat Model with Self-tuning Optimal Regulation of Body Temperature using Hebbian Feedback Covariance Learning", **Medical Physics, (2005),** Vol. 32, No: 12, Pp. 3819--3831.
Ng, E. Y. K. and Ng, W.K. "Parametric Optimisation of the Biopotential Equation for Breast Tumour Idendification using ANOVA and Taguchi Method", **Medical & Biological Engineering & Computing**, Vol. 44, No. 1/2, Springer, Berlin. **(2006)**, Pp. 131-139.
Ng, E. Y-K and Ooi, E.H., "Ocular Surface Temperature: A 3D FEM Prediction Using Bioheat Equation", **Computers in Biology and Medicine**, Elsevier, **(2007),** Vol. 37, No: 6, Pp. 829 -- 835.
Ng, E. Y-K, et al., 2008, The use of tissue electrical characteristics for breast cancer detection: A perspective review, **Technology in Cancer Research and Treatment**, 7(4), 295-308 *(High SCI)*
Ng, E. Y-K., "A Review of Thermography as Promising Non-invasive Detection Modality for Breast Tumour", **International Journal of Thermal Sciences**, Elsevier, Vol. 48, No. 5, **(2009),** pp. 849-855.
Ng, E. Y-K, et al., "Prediction and Parametric Analysis of Thermal Profiles within Heated Human Skin using Boundary Element Method", **Philosophical Transactions A, The Royal Society,** Vol. 368, **(2010)**, Pp. 655--678.
Ng, E. Y-K, and M. Jamil, "Parametric sensitivity analysis of radiofrequency ablation with efficient experimental design: Combined effects of parameters and implications for various tissues", **International Journal of Thermal Sciences**, Vol. 80, No. June, **(2014),** pp. 41-47.
M. Jamil, **Ng, E. Y-K,** "Statistical Modeling of Electrode Based Thermal Therapy with Taguchi Based Multiple Regression", **International Journal of Thermal Sciences**, Vol. 71, **(2013),** pp. 283-291.
M. Jamil, **Ng, E. Y-K,** "To optimize the Efficacy of Bioheat Transfer in Electrode Based Thermal Therapy: A Physical Perspective", **International Journal of Thermal Biology,** Vol. 38, **(2013),** pp. 272-279.

M. Jamil, **Ng, E. Y-K,** “Evaluation of Meshless Radial Basis Collocation Method (RBCM) for Heterogeneous Conduction and Simulation of Temperature Inside the Biological Tissues”, **International Journal of Thermal Sciences,** Vol. 68, (2013), pp. 42-52.
M. Jamil, **Ng, E. Y-K,** “Ranking of Parameters in Bioheat Transfer for Electromagnetic Heating Using Taguchi Analysis”, **International Journal of Thermal Sciences,** Vol. 63, **(2013),** pp. 15-21.

On-going Postgraduate students

No	Student name (funding source)	PhD	Date joined (mm/yr)	Thesis Title
1	(IGS) Suriyanto	PhD	5 Jan 2014	Analysis of renography for quantitative means in differentiating renal obstruction
2	Alan Koh Fuhai, IPP with Lloyd’s Register	PhD	5 Jan 2014	Probability model for ignition of gas ingested by a gas turbine
3	Low Chee Meng (LR, EDB-IPP)	PhD	Aug 2014	Mooring
4	Saxena Ashish	PhD	Jan 2016	Tracking vital health parameters using IR thermal imaging
5	Koh En Wei	MEng	Aug 2016	Decision support system for retinal health using digital fundus images
6	Muhammad Adam Bin Abdul Rahim	MEng	Aug 2016	Application of Thermography for the Detection of Diabetic Foot
7	Oh Shu Lih	MEng	Aug 2017	Application of Deep Convolutional Neural Network for the Identification of Sleep Apnea

Other Research Activities/Achievements (not captured above)

1	Chairman for 15th Int. Conf. on Mechanics in Medicine and Biology (ICMMB-15th, Singapore, 6-8 Dec. 2006)
2	Co-Convenor of the working group on thermal imagers under Medical Technology Standards Committee by SPRING, S’pore, (handling the inter. standardisation on ISO/TR 13154 & IEC 80601-2-59), 2006-2013
3	The "Most Innovative Speaker" Award for 3rd International InfraMatrix Conference & Workshop, “Numerical Predictions of the Ocular Temperature as Alternative to IR Thermography”, KL, Malaysia. (8-10 Nov. 2006).
4	Invited to conduct 3-days lectures in Heat Transfer, Thermodynamic and Mathematical Modeling to Masters (Medical Physics) and PhD students for Dept of Biomedical Imaging, U. of Malaya (2008).
5	Consulted during the on-going swine flu A-H1N1crisis for the setting up of thermal scanner on 1st-line fever check by: Spring-S’pore; ISO/TR 13154:2009; Health Products Regulation Group of HAS; TTSH A&E; Madrid Hospital Foundation of Spain; Centre for Evidence-Based Medicine of Oxford University; STRIDE, Ministry of Defence of Malaysia etc.
6	Best Paper Award for the IRC-SET conference on Science, Engineering and Technology by International Researchers Club. Title: “When Siri Knows How You Feel: Application of Machine Learning in Automatic Sentiment Recognition from Human Speech”, Aug. 2017

Postgraduate students trained to date (Masters or Ph.D. level):

19 Masters and 17 PhD respectively.

No	Student name	PhD	Date graduated (mm/yr)	PhD Thesis Title
1	N. M. Sudharsan	PhD	2000	Numerical modeling of female breast
2	Tan Swee Tiong	PhD	2006	Numerical studies of 3D developing laminar flow in microchannel
3	Zhong Liang	PhD	2006	Biomechanical Engineering Indices for Cardiac Function & Dysfunction during Filling and Ejection Phases
4	Ooi Ean Hin	PhD	2009	Studies of ocular heat transfer using the boundary element method
5	S. Vinithasree	PhD	2010	Early detection of breast cancer using bio-potential field technique
6	Tan Jen Hong	PhD	2010	Development of computer methods in the investigations of ocular surface temperature
7	Seyed Saeid Khalafvand	PhD	2013	Combined CFD/MRI analysis of left ventricular flow in heart failure
8	Muhammad Jamil	PhD	2014	Modeling of cell damage for cancer treatment
9	Jiao Lishi	PhD	2014	Study of Femtosecond Laser Pulse Drilling
10	Koh Wei Xiang Martin	PhD	2016	Complex Interactions between Tidal Turbine System, Multiple

	(ERIAN-IGS, JIP)			Wakes and Seabed Terrain in Energy Capture Array and its associated Environment Issues
11	Chew Kok Hon (DNV, EDB-IPP)	PhD	2017	Optimal Structural Design for Ocean Renewables (Wind) Energy System
12	Chow Jeng Hei (EDB-IPP-DHI)	PhD	2017	Coupled CFD and depth integrated modelling of marine structures
13	Koh Jian Hao (NPGS, JIP)	PhD	2017	Wave loads on floating platforms under special and complex conditions
14	Abdulqadir Aziz Singapore Wala (LR, EDB-IPP)	PhD	2017	Aerodynamics Modelling of Floating Offshore Wind Turbines
15	Ijaz Fazil Syed Ahmed Kabir (RSS)	PhD	2017	Improvement of BEM analysis to incorporate stall delay effect and the study of atmospheric boundary layer effect on the wake characteristics of nrel phase VI Turbine
16	Tejas Canchi, (LKCMedicine's RSS)	PhD	2017	Mechanistic and Pathological Study of the Genesis, Growth, and Rupture of Abdominal Aortic Aneurysms
17	Nikhil Garg	PhD	2017	The Effects of Atmosphere-Ocean-Wave Coupling During Tropical Cyclone

Number of post-doctorates and PhD students currently in the lab:
Nil for PDF.

The recently completed externally funded project "Discrete Thermal Data Analysis using ANN" by Lifeline Biotechnologies, Inc. (# RCA2-NTU, though <$500k) has granted 3 patents (with Lifeline Biotechnologies , Inc. NV, received USD10m of investment in USA) and related good journal papers as listed below:

Method/Process: "A Device for Analyzing Thermal Data Based on Breast Surface Temperature for the Detection for Use in Determining Cancerous Conditions".	US Patent No: 8,185,485 B2
System Utility: "A System for Analyzing Thermal Data Based on Breast Surface Temperature to Determine Suspect Conditions"	US Patent No: 8,231,542 B2
Method Utility: "Methods for collecting and Analyzing Thermal Data Based on Breast Surface Temperature to Determine Suspect Conditions"	US Patent No: 8,226,572 B2

1. M. EtehadTavakol, S. Sadri, Ng, E.Y.K, 2010, Application of K- and Fuzzy c-means for Color Segmentation of Thermal Infrared Breast Images, Journal of Medical Systems, 34(1):35-42. (IF=1.064)
2. Tan, JMY, Ng, E.Y.K., R Acharya U, Keith, L.G., and Holmes, J., 2008, Comparative Study on the use of Analytical Software to Identify the Different Stages of Breast Cancer using Discrete Temperature Data, Journal of Medical Systems. 33(2):141-153 (IF=1.064)
3. Ng E.Y.K., et al., 2007, Detection and Classification of Breast Cancer using Neural Classifiers with First Warning Thermal Sensors, Information Sciences, 177(20):4526-4538. (IF: 3.095)
4. Tan & Ng E.Y.K. et al, 2007, A Novel Cognitive Interpretation of Breast Cancer Thermography with Complementary Learning Fuzzy Neural Memory Structure, Expert System With Applications, 33(3): 652-666. (IF= 2.596)

Another 2 externally funded projects "Biopotential field detection of breast cancer" and "Thermal imaging of eye diseases" by Tote-Board-Ngee Ann Gong-Xi (with TTSH & NUH respectively, though <$500k) have also resulted some related good journal papers as listed below:

1. Tan JH, Ng E.Y.K, Rajendra Udyavara Acharya, Chee C, 2009, Infrared Thermography on Ocular Surface Temperature: A Review, Infrared Physics & Technology, 52(4):97-108. (IF=1.037)
2. Tan & Ng et al, 2010, Evaluation of Tear Evaporation from Ocular Surface by Functional Infrared Thermography, Medical Physics, 37(11):6022-6034. (IF=3.871)
3. Ooi EH, Ng EY.K., 2008, Simulation of Aqueous Humor Hydrodynamics in Human Eye Heat Transfer, Computers in Biology and Medicine, 38(2):252-262. (IF: 1.272)
4. Ng Y.K., S Vinitha Sree, K H Ng and G. Kaw, 2008, *The use of tissue electrical characteristics for breast cancer detection: A perspective review*, Technology in Cancer Research and Treatment, Vol. 7, No. 4, pp. 295-308.
5. Ng Y.K., Ng Wan Kee, Sim SJL, Rajendra Acharya U, 2007, *Numerical Modelling of Biofield Potential for Detection of Breast Tumor*, Computers in Biology and Medicine, Vol. 37, No. 8, pp. 1121-1132. (Q1 J.)

Chapter 1

Stage 1:
Idea Generation

Think and think wild. Keep a curious heart.

Initial Project

Initial Project Title: Organization of Medical Data as an Input for Machine Learning Processes

Initial aim of the project: A computational method (machine learning) is being used to develop a diagnostic tool for a vascular disease. This will have implications for clinicians to better diagnose the patient's condition so that suitable surgical care can be delivered.

Procedure: As part of a research project, a large amount of patient data will be available in the form of various parameters of each patient. This may be up to 30 in number for each patient. It is envisioned to have about 200 patients worth of data as input to the machine learning method in the proper format. It is important for the data to be input in a particular format, so that the diagnostic tool may predict the statistical distribution accurately.

Requirement: There is a need for a data entry specialist to manage the large amount of data that is available through the concerned hospital.
Responsibilities: The person needs to accurately manage the various parameters of each patient. Presentation and maintenance of data in the prescribed format so that it may be used for the machine learning method.

The person must be familiar / proficient with tools such as MS Excel or any other data management system.

New Project Idea Generation 1

NRP Project brainstorm

Core Method: Machine Learning

Application: Language

Possible direction	My interest	My ability	Relevance to my society	Sourcing data
Accent Recognition				
Text genres categorization				
Cultural context identification				
Information extraction				
Sentiment analysis				
Handwriting recognition				

Using this “Research Interest Matrix”, I decided to go to the field of Sentiment Analysis, although I have absolutely no background in it and the field is different from my supervisor’s speciality.

Conclusion: Sentiment analysis

- Automated text classification for spoken texts / written texts based on attitudes using machine learning

Methodology:

Corpora + machine learning

Spoken? Written?

Cultural?

Textual?

Experiential?

Interpersonal?

Emotions? Opinions?

Whatsapp?

Features	Aim
Voice level (Pitch, speed, amplitude, fluctuation) Textual level (Pauses, connotation of words, power, pronouns)	Predict emotions
Connotation of expressions (words, phrases, idioms), semantic field,	Classify opinions

New Project Idea Generation 2

<table>
<tr><th colspan="3">Project idea 1</th></tr>
<tr><td>Title:</td><td colspan="2">Emotion recognition from speech using machine learning</td></tr>
<tr><td>Description:</td><td colspan="2">Train machine to recognise emotion in speech on a sentence level, using speech features including acoustic cues and linguistic features. In this project, emotions are categorised into anger, boredom, disgust, fear, happiness, sadness and neutral state.</td></tr>
<tr><td>Application:</td><td colspan="2">Higher level human-machine interaction</td></tr>
<tr><td rowspan="2">Data:</td><td colspan="2">Speech occurred in converstaions</td></tr>
<tr><td>Audio</td><td>Transcript</td></tr>
<tr><td>Feature extraction:</td><td>Acoustic features:
• Pitch
• Intensity
• Jitter</td><td>Textual features:
• Connotation of words
• Emotional words
• Phrase length</td></tr>
<tr><td>Pre-processing of data:</td><td>• Sound processing software
• Normalisation (noise factors such as gender differences)</td><td>• Natural Language Processing
○ Existing parsing tools
○ Corpora
○ Online word databases</td></tr>
<tr><td>Labeling of training data:</td><td colspan="2">Manually label the emotion corresponding to each speech</td></tr>
<tr><td>Machine Training:</td><td colspan="2">Choose the right machine learning method</td></tr>
<tr><td>Output:</td><td colspan="2">Emotion of the given speech</td></tr>
<tr><td>Potential problems:</td><td colspan="2">1. Must I use original data? Can I use data from the Internet?
2. Can I use speeches extracted from fictional movies/theatre shows?
3. Will this be possible/feasible?</td></tr>
</table>

<table>
<tr><th colspan="2">Project idea 2</th></tr>
<tr><td>Title:</td><td>Opinion mining using machine learning</td></tr>
<tr><td>Description:</td><td>Use news articles as training data to train machine to analyze the text opinion. In this project, opinions are categorized into neutral, negative (against), and positive (for).</td></tr>
<tr><td>Application:</td><td>Analyzing business review, movie review, political opinion, so as to make predictions/decisions; mining opinions of historical events and current affairs.</td></tr>
<tr><td>Data:</td><td>News commentaries from wide range of prominent and popular sources, for example, The Straits Time, The Economist, The New York Times etc.</td></tr>
<tr><td>Feature extraction:</td><td>• Natural Language Processing
○ Existing parsing tools
○ Corpora
○ Online word databases</td></tr>
<tr><td>Labeling of training data:</td><td>Manually label the stand of each articles/opinions</td></tr>
<tr><td>Machine Training:</td><td>Choose the right machine learning method</td></tr>
<tr><td>Output:</td><td>Stand of the given article</td></tr>
<tr><td>Potential problems:</td><td>1. Access to some of the news archives</td></tr>
</table>

Comparison	Project 1	Project 2
Collection of data	Reuse data? Start recoding conversations from scratch? Extract recordings from movies/ shows? Quite difficult.	Only need to get the access to the news archive
Processing data / feature extraction	More difficult	Only need to worry about natural language processing techniques
Labelling	Manually	Manually
Innovativeness	Relatively new, more exciting prospect	Done before, even commercialized in some sectors

Finalized Project Idea

Finalized Project Idea: Project Idea 1

Title:

Automatic emotion recognition from human speech using machine learning

Description:

Background: Machine Learning (ML) is a vibrant research area in Artificial Intelligence (AI) and can be applied to a diverse range of fields. One of the areas is Human-Machine Interaction (HMI) which, at a higher level, may involve communication and understanding of emotions.

Aim: This project explores the possibility of using a machine learning approach to recognize emotions from utterances (on a sentence level) in English language. Both acoustic features and textual features will be considered in order to increase the accuracy of automatic emotion recognition.

Methodology:

Raw data: Online speech databases which provide recordings of emotional utterances.

Feature extraction: Extract both acoustic features (including pitch, intensity, and jitter), using sound processing software and textual features (including word connotations, detection of emotionally salient words, phrase length), using natural language processing tools.

Pre-processing of data: Normalize the extracted features and weigh the saliency of different features in contributing to emotions. Organize the data in an appropriate format for input, so that the machine can predict the statistical distribution accurately.

Machine learning process: Input training data to appropriate classifiers so as to achieve optimization.

Chapter 2

Stage 2:
Literature Review

Stay sane. Have an inquisitive mind.

Section 1

Learning Materials

Python Machine Learning (Raschka, 2015)

Application of Machine Learning Methods in Assessment AAA Rupture Risk (Yong Yam Yuan, 2016)

Sentiment Analysis and Opinion Mining (Bing Liu, 2015)

Computational Paralinguistics (Bjo¨rn W. Schuller &

Anton M. Batliner, 2014)

Affective Computing (Rosalind W. Picard, 1997)

NLP course by Michael Collins, Columbia University

NLP course by Dragomir R. Radev, Ph.D, University of Michigan (https://www.coursera.org/learn/natural-language-processing)

Part of Machine Learning course by Andrew Ng, Stanford University

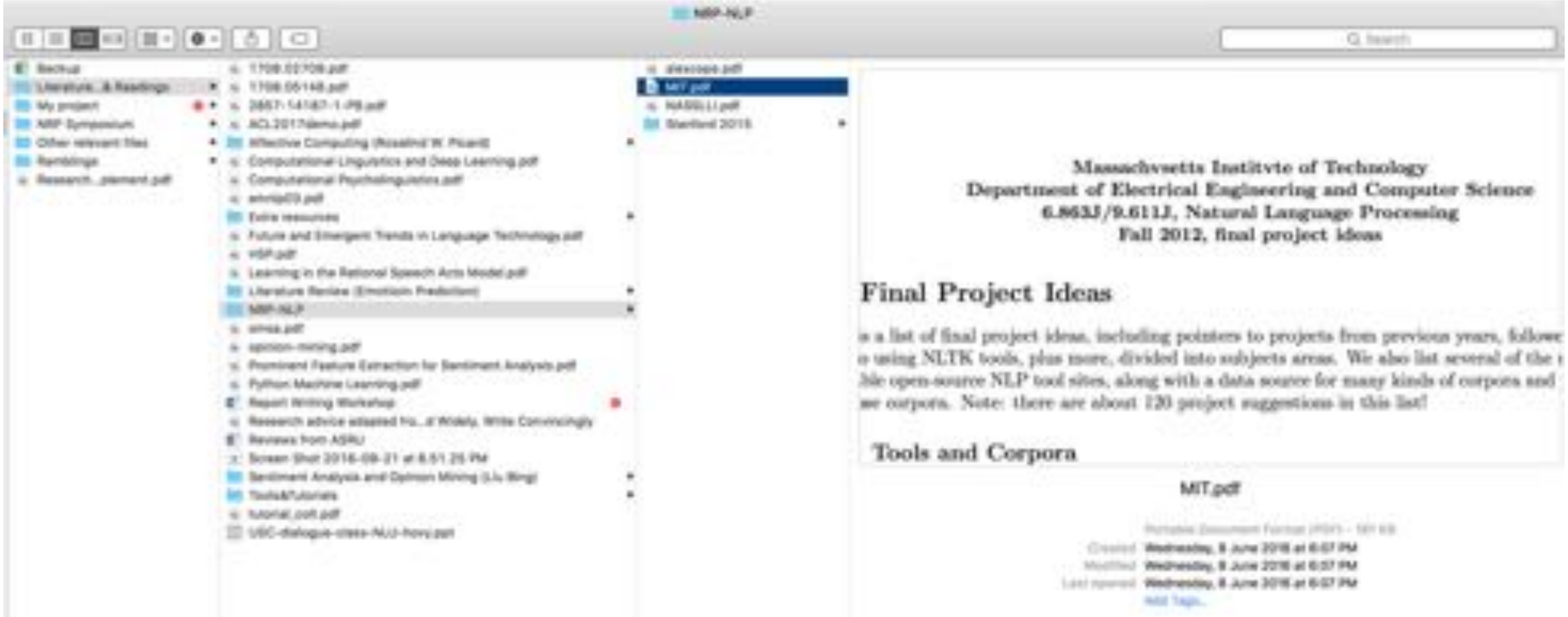
Part of Artificial Intelligence course by Dan Klein &

Pieter Abbeel, University of California, Berkeley (https://courses.edx.org/courses/BerkeleyX/CS188x_1/1T2013/info)

Pattern Recognition and Machine Learning by Christopher Bishop

Section 2

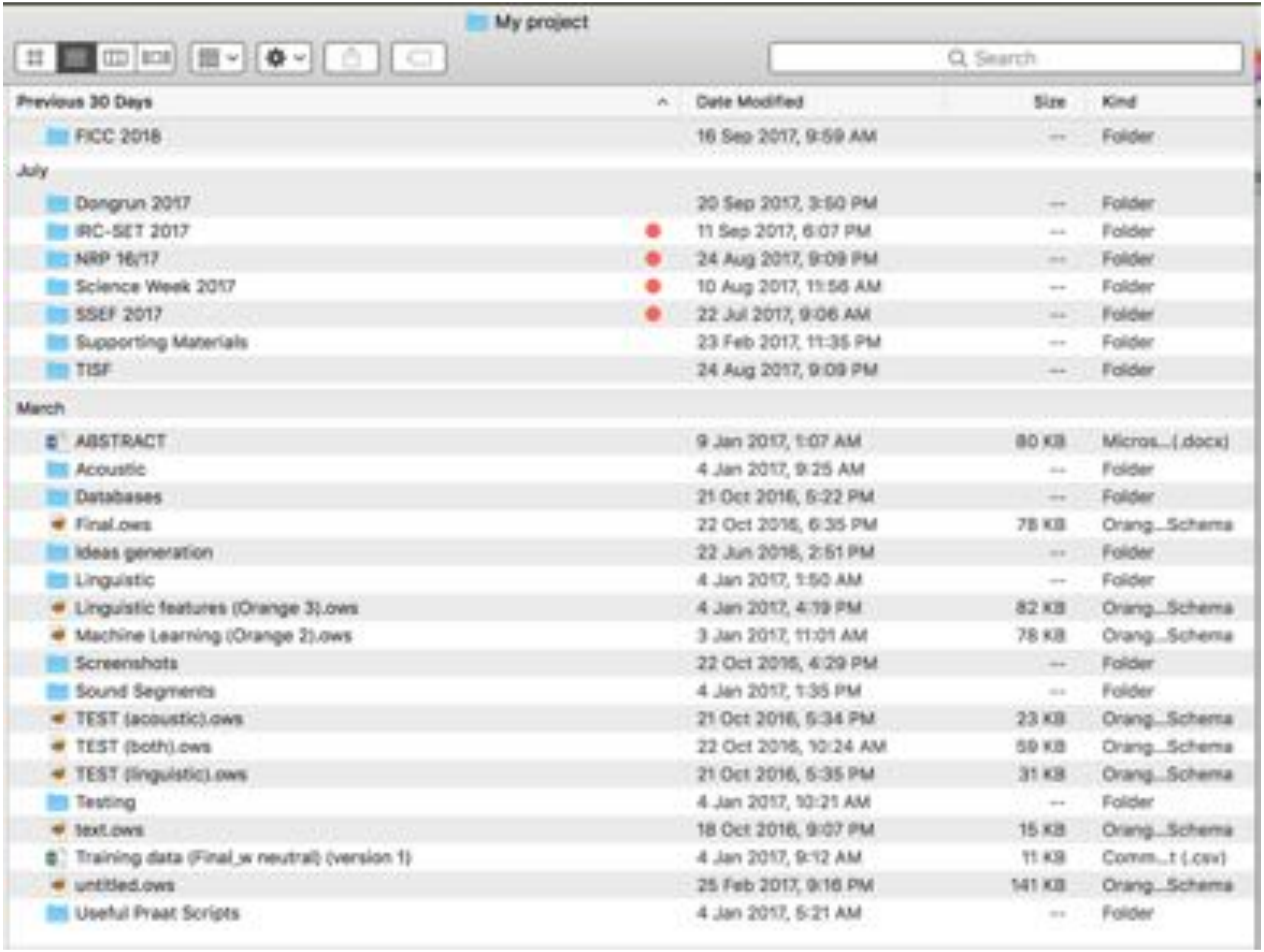
Related Papers



The above shows the folder that I use to keep the papers that I have read, which are simply too many to be all listed here.

Since I was entirely new to the field, I had to read papers as widely as possible. Therefore, initially I read many seemingly irrelevant papers and research. They were useful nonetheless as a comprehensive and intriguing introduction to the field of Natural Language Processing. Many papers provided very interesting and challenging direction of studies and I was informed about the enormous diversity and possibility in the field

The following is a screenshot of my project files.



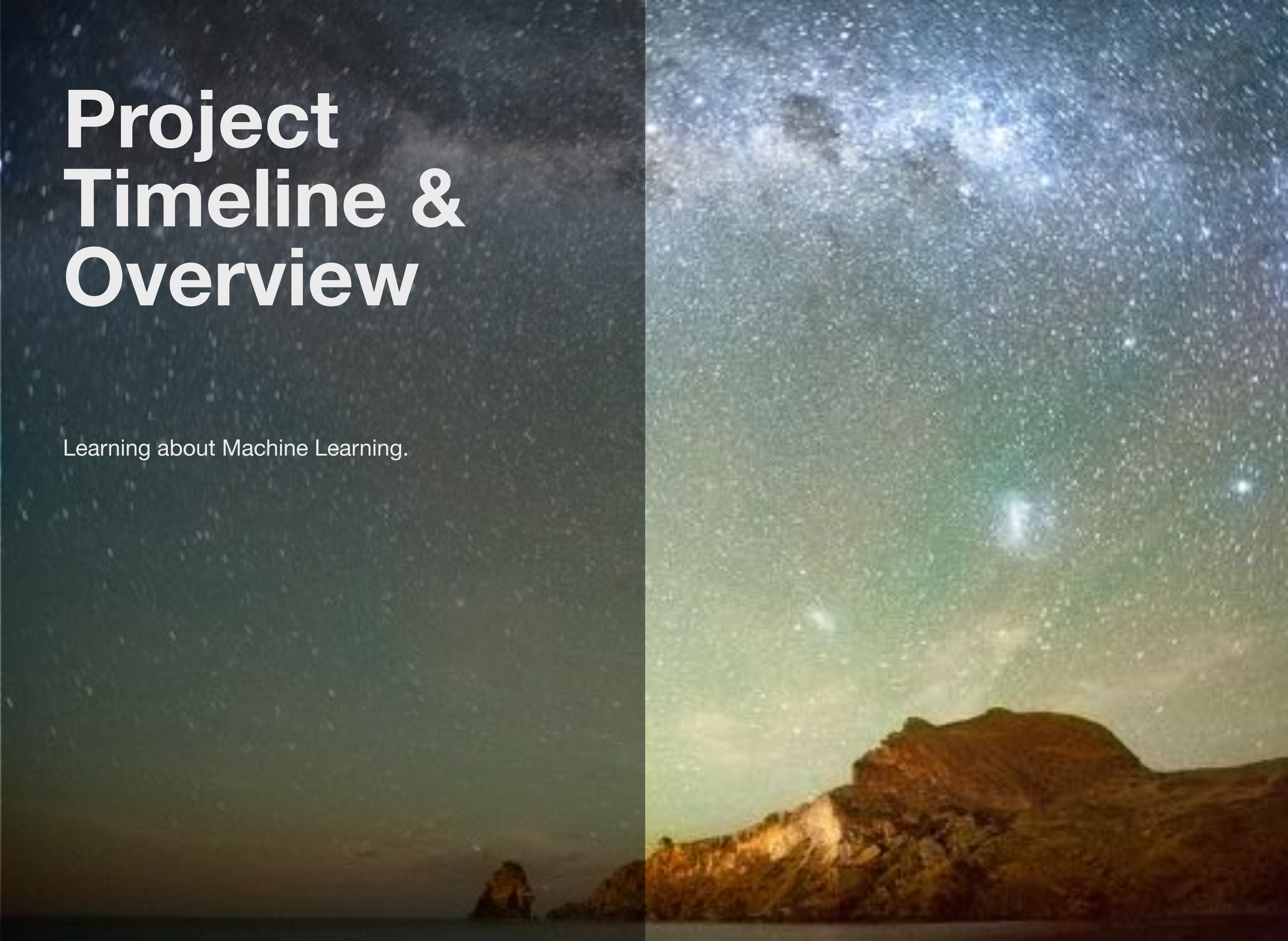
My project

Name	Date Modified	Size	Kind
Previous 30 Days			
FICC 2018	16 Sep 2017, 9:59 AM	--	Folder
July			
Dongrun 2017	20 Sep 2017, 3:50 PM	--	Folder
IRC-SET 2017	11 Sep 2017, 6:07 PM	--	Folder
NRP 16/17	24 Aug 2017, 9:09 PM	--	Folder
Science Week 2017	10 Aug 2017, 11:56 AM	--	Folder
SSEF 2017	22 Jul 2017, 9:06 AM	--	Folder
Supporting Materials	23 Feb 2017, 11:35 PM	--	Folder
TISF	24 Aug 2017, 9:09 PM	--	Folder
March			
ABSTRACT	9 Jan 2017, 1:07 AM	80 KB	Micros...(.docx)
Acoustic	4 Jan 2017, 9:25 AM	--	Folder
Databases	21 Oct 2016, 6:22 PM	--	Folder
Final.ows	22 Oct 2016, 6:35 PM	78 KB	Orang...Schema
Ideas generation	22 Jun 2016, 2:51 PM	--	Folder
Linguistic	4 Jan 2017, 1:50 AM	--	Folder
Linguistic features (Orange 3).ows	4 Jan 2017, 4:19 PM	82 KB	Orang...Schema
Machine Learning (Orange 2).ows	3 Jan 2017, 11:01 AM	78 KB	Orang...Schema
Screenshots	22 Oct 2016, 4:29 PM	--	Folder
Sound Segments	4 Jan 2017, 1:35 PM	--	Folder
TEST (acoustic).ows	21 Oct 2016, 5:34 PM	23 KB	Orang...Schema
TEST (both).ows	22 Oct 2016, 10:24 AM	59 KB	Orang...Schema
TEST (linguistic).ows	21 Oct 2016, 5:35 PM	31 KB	Orang...Schema
Testing	4 Jan 2017, 10:21 AM	--	Folder
text.ows	18 Oct 2016, 9:07 PM	15 KB	Orang...Schema
Training data (Final_w neutral) (version 1)	4 Jan 2017, 9:12 AM	11 KB	Comm...t (.csv)
untitled.ows	25 Feb 2017, 9:16 PM	141 KB	Orang...Schema
Useful Praat Scripts	4 Jan 2017, 5:21 AM	--	Folder

Chapter 3

Stage 3:
Experimentation

Trial and error. Be a brave soul.



Project Timeline & Overview

Learning about Machine Learning.

Timeline

Date	Event
21 Apr 2016	1st Meeting (Introduction)
Apr – May	- Download Orange Canvas and try tutorials - Learn basic syntax in English language
9 May 2016	2nd Meeting (Ms Yong's Final Year Project)
20 May 2016	ACJC Scientific Literacy Workshop by A*STAR
Jun	− Self-learning [Relevant parts only] ○ The book *Python Machine Learning (Raschka, 2015)* ○ Coursera: *Machine Learning (Stanford University)* and *Natural Language Processing (Columbia University)*
Aug –Sept	- Finish Progress Report - Self-learning ○ Coursera: *Introduction to Natural Language Processing (University of Michigan) https://www.coursera.org/learn/natural-language-processing* ○ EdX: *Artificial Intelligence (University of California, Berkeley)* https://courses.edx.org/courses/BerkeleyX/CS188x_1/1T2013/info

Date	Event
By Oct 2016	Analysis of data (after acquiring the access to database)
Oct 2016	- Examine the existing language technologies and compare the theories behind them - Continue the online courses and build theoretical foundation, preferably also the Math behind - Test results, record, analyze, and improve
Nov 2016	- Targeted reading to solve technical challenges met in the previous experiments - Make enhancements to previous methodology - Self-learning: Sentiment Analysis and Opinion Mining (Liu Bing)
Dec 2016	Conclude and write the Research Paper
8 Jan 2017	Submit the Final Report
Feb - Mar	Make project poster
8, 9 Mar	SSEF Finals
6 Apr	NRP Symposium; Project Showcase
6 Aug	IRC-SET Paper Presentation and Poster Showcase
5 Nov	Dongrun-Yau Computer Science Award Semi-Finals

Section 2

Overview

What is the eventual shape of my project? Let's have a glance through my final abstract:

Opinions and sentiments are essential to human activities and have a wide variety of applications. As many decision makers turn to social media due to large volume of opinion data available, efficient and accurate sentiment analysis is necessary to extract those data. Hence, text sentiment analysis has recently become a popular field and has attracted many researchers. However, extracting sentiments from audio speech remains a challenge. This project explored the possibility of applying supervised Machine Learning in recognizing sentiments in English utterances on a sentence level. In addition, the project also aimed to examine the effect of combining acoustic and linguistic features on classification accuracy. Six audio tracks were randomly selected to be training data from 40 YouTube videos (monologue) with strong presence of sentiments. Speakers expressed sentiments towards products, films, or political events. These sentiments were manually labelled as negative and positive based on independent judgement of 3 experimenters. A wide range of acoustic and linguistic features were then analyzed and extracted using sound editing and text mining tools respectively. A novel approach was proposed, which used a simplified sentiment score to integrate linguistic features and estimate sentiment valence. This approach improved negation analysis and hence increased overall accuracy. Results showed that when both linguistic and acoustic features were used, accuracy of sentiment recognition improved significantly, and that excellent prediction was achieved when the four classifiers were trained respectively, namely kNN, SVM, Neural Network, and Naïve Bayes. Possible sources of error and inherent challenges of audio sentiment analysis were discussed to provide potential directions for future research.

Chapter 4

Retrospects

Better and better. Never stop at the finishing line.

My notes and thoughts while/after doing the project.

On The Limitations of Current NLP Research (my own thoughts)

Natural **'discourse'** processing

Much of the current NLP research seems to over-simplify language. Of course, it is understandable because we want to use computational and mathematical symbols to represent natural language so that it is able to be processed - 'understandable' - by the machines. However, one fundamental problem with this approach is that the complexity of the interactions between language and speakers, contexts are overlooked. This would have far-reaching implications, especially in NLP applications such as dialogue systems.

Language is meaningless without CONTEXT; it does not happen in a vacuum. However, most of the NLP applications treat language as something standing on its own. Granted the high accuracy of advanced parsing techniques and advanced classification methods such as HMM. What we seem to neglect all along is the context where all these linguistic data are being produced. For example, the world's most advanced parser will fail to analyze the complex meaning of the seemingly simple utterance 'I am good'. The speaker might be replying casually to a greeting. The speaker might have a bad day and do not want to be disturbed.

In short, NLP, especially higher level NLP, has to take into consideration not only the quantitative, but more importantly, the qualitative aspect of things. Yes, but we cannot and should not simply throw linguistic data to the machine and treat any linguistic data indiscriminately. In order to achieve real NLP, we have to incorporate the current research in linguistic and language. And that requires collaboration between linguists and computer scientists to work together. That requires us to have a brand new mindset towards NLP.

New Project Ideas Brainstorming (my own thoughts)

Parsing techniques for chemical compounds

L1: Automatic classification of compounds

L2: Automatic prediction of products of reactions

L3: Chemistry-inspired model for analysis of language change and language variation

L4: Economic policy assessment using

Design an AI that receives an external signal through sensors, such as smelling something, identify it, and translate it into natural language for users to understand.

Section 3

My New Project Idea: Language Evolution and AI

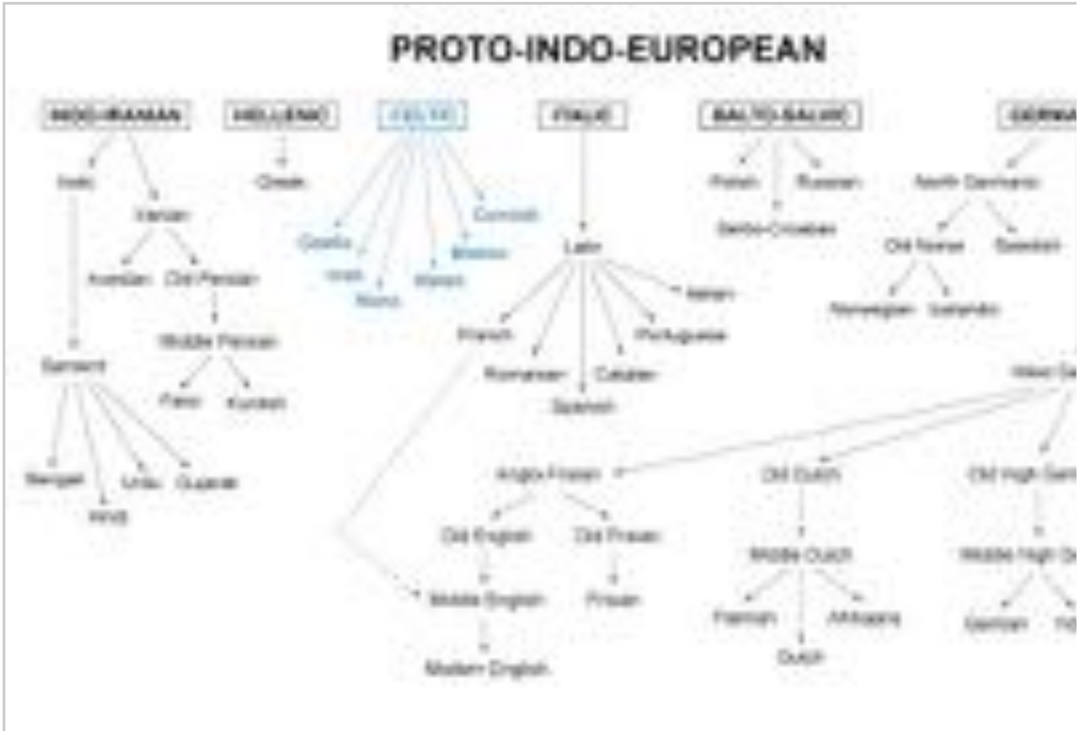
Language contact as a chemical reaction.

The higher the prestige, the higher the stability --> always approaching lower energy

<u>Language Entropy Hypothesis</u>

Given sufficient time, a variety of language will eventually and definitely becomes a language on its own.

The other day I had a wild guess (you could call it a hypothesis/postulate) about natural language and artificial language. Since the AI learning algorithm is modelled after human intelligence, we could draw some insights as to how humans have developed our natural language. As shown below, one language branches out to form ‘new language varieties’ due to language contact. And then eventually these varieties of language become new languages when there are sufficient differences from the origin language. Probably co-incidentally, we observe the same trend on the Asian continent - ancient Chinese spread to and eventually evolved into Japanese and Korean languages.



So my hypothesis is that **give sufficient time**, a variety of language will **eventually and definitely** becomes a language on its own. This hypothesis, if proven right, could be extended to artificial language (the language people purposely create) and AI language (the new variety that AI develops). (probably programming language as well? I don't know.)

After having a vague idea, I had a conversation with a chemistry teacher. He actually used the chemical equilibrium and thermodynamics model to predict that in the future there might be only one universal language in the world. This is because like a reaction system, the complexity of language is initially low (only one original porto-indo-european language) and then increases (many different languages is developed) and eventually approaches homogeneity (maximum randomness), when it tends towards a single language that is like a 'rojak'. [this is disputable but even some researchers also have similar ideas, out of different reasons]

Conversation with Drama Teacher, Mrs Creffield

1. Introduce my AI project briefly

2. Especially highlight the potential applications

3. Discuss on the limitations and future directions, especially those related to emotions

4. How do we deal with the complexity of human emotions?

5. If there can be a set of rules for actors to follow in order to break down and construct emotions, can we use them to design affective/sentient AI as well?

Subtext: withheld

Silence intentional understating

Humble bragging: evasion, understatement

Intentional understating

Hyperbole: overstate: disguise uncertainty

The speech gap

Inflection cadence

ESL,EFL & native speakers

Origin of the speech

Context power dynamics

Who and why are you addressing?

Social-cultural frame

Paraliguistics:

1. Vicarious Body

2. Corperial body

3. Kinesthetics

4. Proxemics: distance

Embodiment

1. Danger of stereotypes

2. Racialization

AI and racial prejudice

Conversation with Chemistry Teacher, Mr Kuna

1. Explain briefly my hypothesis about language change

2. How could we possibly make sense of language change through chemical model?

3. How should we approach such a hypothesis? How do we prove it? Can we use Chemical Modelling?

4. Are there any limitations? How do we overcome them?

Adaptive radiation

Speciation

Natural selection to give it a pressure

Language evolution: a trait being selected for and against

Some degree of order initially

A degree of complexity

Tends to homogeneity

Borrowing or creating new words

Originality is lost morphs into something new

Describing patterns -->> model --> and apply to other situation extrapolation

Measurements involved

Science is a method, a way of knowing

Observation **(pattern): data**

Formulate hypothesis

Test out the hypothesis: is it only one-off or can apply to other situations?

Theory / model

Predictions (and to some extent control and manipulate)

If such prediction is born out: if the pattern is repeated

Language subjective elements: see across different languages

Science behind how emotions are constructed

Model is there a **universal** e.g. Ma, mother, ah ma: if there is a scientific underpinning to it

See the parallel

Logic

Mutual impact, action and reaction

Semblance model in some aspects

Urdu = hindi + tamil

Practiced differently even under the same name

Increasing homogenising (underlying law for everything)

everything is driven by gradient

Scientific basis in language

Quantum law of complementarity: 波粒二象性

Wisdom revealed by scientific truths:

E.g. ‘Dissolving’: cations being seduced by the water molecules

‘faithful’

Laws are aggregate behavior

A lot of randomness in individuals / atoms

Break symmetry create function: hybridisation

Create the spherical symmetry to form directional orbital

Language preservation is to fight against entropy.

More about “Entropy”:

1. Charge delocalisation => stress on people

Charge density high: unstable => people with too much stress are emotionally unstable

Charge dispersion => share out stress => physics spread out pressure

2. Equilibrium is death: no more activity

3. Everything tends towards a state of equilibrium
--> max entropy

People absorb energy --> within themselves, more orderly; but create more chaos outside --> and the increase in chaos >> increase in orderliness --> eventually, the world tends to a state that has the highest entropy

4. From physics snell’s law to reaction pathway
--> min energy

5. Pythagoras theorem and $\sin^2 + \cos^2 = 1$

6. Chem and econs basically the same principles/ concepts but just different terminology

Computional Imagining Problems (from SMART researcher)

AI

Light -> film

Light -> retina -> brain

Light -> CMOS -> CPU

“Phase” object cannot be seen

Except indirectly

- defocus
- Zoom
- Dust
- Reflection

Spatial map of phase

A phase odyssey

Inverse problem:

Tikhonov-Wiener: fitness term + regularizer

H and must be known or discoverable

Shen Zuowei

Digital holography

Caustics (singularities)

1. General form

G = HF

2. Tikhonov-Wiener

3. New method:

G -> neural network -> f^

Both (2&3) use optimisation

Classification vs inversion

Undertrained --> overtrained (curve similar to MES EOS curve)

Overtrained --> too generalised

The good thing about NN: when it fails, it fails gracefully.

Deep, deep trouble

SLM, faces-LFW, ImageNet