HeartRateStress: Can Heart Rate features detect subjective stress?

Pradyoth Hegde, Santhosh Subramanian and Tianju Ma

***Abstract*— Stress is defined as a feeling of mental or emotional strain resulting from different circumstances. This strain can lead to several diseases including but not limited to cardiovascular disease, diabetes, asthma and other chronic illness. New fitness trackers claim they go beyond counting steps and measuring sleep. Currently, these devices can also identify stress. The goal is to help people identify the factors which trigger stress, and try to steer clear of them. The primary objective of this project is to predict subjective stress based on the heart-rate features.**

In this paper, we built a model that can predict stress in real time. The model was trained using the data collected from ten participants in a lab study. The participants were asked to wear a LG Urbane smartwatch, a commercial product, and were told to perform different activities which induce stress. We extracted several statistical features over a window size of one second from the obtained data and built a Random Forest model that can predict stress with an accuracy of 90.12 %. Based on the data collected from each activity and also from the self-reported stress values, sing-a-song was the most stressful activity.

I. INTRODUCTION

The first term to be defined through the course of this paper is Stress. Since it is a such a subjective phenomena, it might be hard to define it. However, the term stress, as it is currently being used was defined by Hans Selye in 1936 as the non-specific response of the body to any demand for change. In his experiments with laboratory animals, which were subjected to acute but different stimuli such as blaring light, deafening noise and extremes of hot or cold temperatures, the animals showed symptoms such as stomach ulcerations, adrenal enlargement etc. He later demonstrated that persistent stress could cause these animals to develop diseases which were similar to humans, such as heart attacks and kidney disease. Selye proposed that different phenomena can cause the same disease and that this was not just limited to animals, but humans as well. [4]

There have been many instances when individuals feel that they are able to work better under certain pressure or stress However increased and prolonged periods of stress is very harmful and can cause different ailments such as depression, fatigue and aggression. A recent study shows that one in every four Americans suffer from Stress [7]. That's roughly 80 million Americans and many more throughout the world. Therefore, stress detection and necessary steps to combat it is definitely a task which would be substantial. There has been a lot of success in passive sensing stress research[1] and using these techniques to build a machine learning model was effective.

The aim of this project was to focus on determining and analyzing the vital signs which show significant change when a person is stressed. In order to do this, two sensors were used. The first one being LG Smart-watch Urbane, which is a commercial smart-watch and the second being "Bio-Stamp" which is a body-worn sensor. We built applications which interface with these devices in order to acquire data. We gathered data from ten participants and sent these values to a back-end server. After extraction of data, we built a model using Machine Learning techniques. Afterwards, we used the parameters obtained from our model and fed them into our application, which was able to detect stress levels in real time. The results obtained are discussed, along with sections for conclusion and future scope.

II. RELATED WORKS

There have been different approaches to detect Stress. StressSense [8] measured stress from human voice using a smart-phone microphone which recognizes changes in a person's speech. However, this had a limitation that the user had to constantly speak which was quite inconvenient. Galvanic skin response has also been used in order to determine stress [9] Vrijkotte et. al. detects work life stress using blood pressure, heart rate and heart rate variability. Since this showed satisfactory results, the concept of using heart rate based features to determine stress seemed promising. Also, to be able to build an Android Application which could predict in real-time, whether a person was stressed or not, seemed like the perfect application of Wireless and Mobile Health Technologies.

There has been a lot of research to determine stress levels of a person using various vital signs. One of them is cStress[1]. The primary purpose of this paper was to obtain a gold standard which was reproducible and had a very high accuracy. They used a sensor suite, which included many biomedical sensors, amongst which the most important ones were a two lead Electrocardiograph, a flexible chest band and a 3-axis accelerometer.They came up with effective methods to process data, such as dynamic programming and cubic Hermite splines to interpolate the gaps. Some of the evaluation methods used in our paper are based from this.

Another important research paper which helped is uStress. [5] This paper mainly focuses on defining predictive features to detect subjective stress in college students from wrist and chest-worn sensors, along with the forms of stressors that are most effective in inducing stress. The paper describes a protocol which served as a guideline, where certain stressful and non stress activities are defined with some amounts of rest period in between them. Various machine learning techniques were used in order to allow them to detect stress in real time. They have concluded that participant singing causes the highest stress levels.

III. System

The objective of this paper is to determine whether heart rate features would be able to detect subjective stress. This was done by inducing stress in a controlled environment and different features were identified, evaluated and analyzed. Participants' response to various forms of stressors were recorded and the model was validated.

A. Devices and Sensors

1) LG Smart-watch Urbane: This is a commercially available smart-watch which has a heart-rate sensor and can provide an accurate measure of heart-rate. It also has a tri-axial accelerometer, a proximity sensor and a gyroscope. The device provides heart-rate values in the intervals of one second. It communicates with the smart-phone via Bluetooth low energy (BLE) and also has Wi-Fi access to send data directly to a back-end server. The smart-watch can be seen in Fig (1)
2) Bio-Stamp: The Bio-stamp sensor is a tattoo based sensor which is very flexible and accurate. It has a tri-axial accelerometer, an ECG, and an EMG sensor. It communicates with a computer via BLE and has a sampling frequency that can be configured based on the requirement. For this application, a sampling frequency of 250Hz was used. The Bio-stamp provides an output of raw ECG values, which had to be processed and analyzed in order to obtain heart rate features. The Bio-stamp sensor can be seen in Fig (2)

Fig. 1. LG Smart-watch Urbane

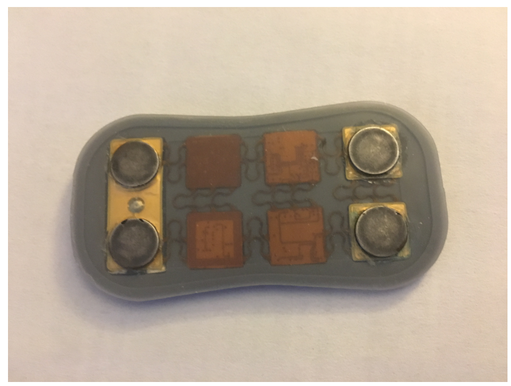

Fig. 2. Bio-Stamp sensor

B. Experiments for Inducing Stress

In these experiments, participants were invited to a lab, and were requested to wear sensors. They did not know what tasks they were about to perform; they wore the sensors and sat in front of a computer. The computer had a presentation which detailed each activity and the duration for which it had to be performed.

Table (1) shows details of the activities performed. The activities were randomized for each participant. Each activity had to be done for four minutes, except the cold pressor test, which the participants could do for as long as they could. Between each of these activities, whether stressful or non stressful, there was a rest period of two minutes. At the end of each stressful activity, the participant was asked to rate their stress level on a scale of 1-10, 1 being the lowest and 10 being the highest. The activities were labeled as Stressed(S), Not-Stressed(NS) and Rest(R).

Fig. 3. Participant performing given protocol

TABLE I
ALL ACTIVITIES, CORRESPONDING LABELS AND DESCRIPTION.

Activity	Label	Description
Game	S	Play the game ”The case of Scary Shadows”
Arithmetic Test	S	Answer as many questions correctly as possible before time runs out
Stroop Test[12]	S	Enter the first letter of the color of the text shown in activity and get as many correct as possible until time runs out
Sing-a-Song[11]	S	Sings songs loudly without any background music
Eating	NS	Enjoy the complimentary food
Social Conversation	NS	Have a conversation with the experimenters
Emails	NS	Check emails and respond to anything if important
Cold Pressor [10]	S	Completely immerse non dominant hand in ice cold water for as long as possible

C. Application Development

In order to extract heart-rate values from a participant, we built an android wear application which ran on the LG Smart-watch. The flowchart of the application is described in Fig (7)

D. Data Collection

In order to collect heart-rate values from participants, an Android Wear Application was developed which stores heart-rate values and corresponding timestamps for a certain time period in a local database. The heart-rate values are calculated every second. When required, it sends this data to a back-end server. The sampling frequency is 1Hz

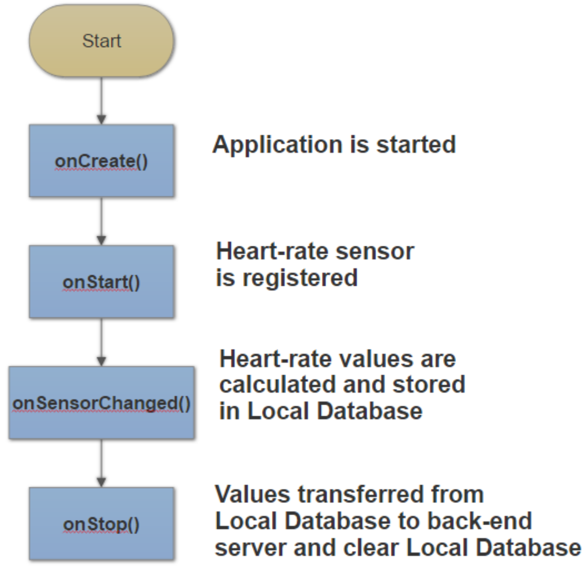

Fig. 4. Flowchart of Application

An example of raw data collected from the Smart watch can be shown in Fig (5). This shows the plot between the heart-rate and time. Additionally, the different activities done by the participant are also marked to provide better understanding. The high peak of the orange line indicates a participant is doing an activity. The low peak indicates a rest period.

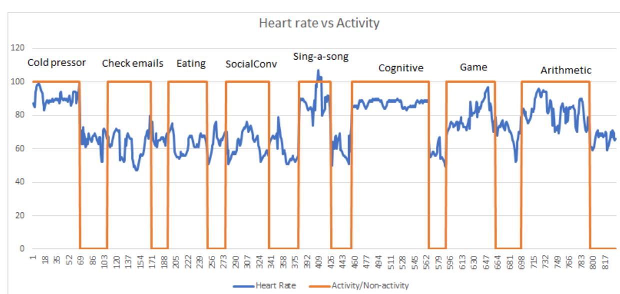

Fig. 5. Raw Data values for one participant

IV. METHODOLOGY

This section talks about the methodologies used to predict subjective stress.

1) Feature Extraction: There were ten participants who performed the given tasks. Features were extracted on each time-series data stream and all the features were based on statistical analysis of the data across the entire activity. A minute based approach was used in order to extract the features. In this approach, we used Passive Sensing Data Analytics Chain (PASDAC) code with a window size of one second and a sliding window protocol to obtain the features as shown in Table (2). Timestamps for different activity periods were noted and were cross referenced with the experimental data, and were classified accordingly in the training set.

TABLE II

FEATURES EXTRACTED

Feature	Description
Mean	Average of values in the window
Max	Maximum value in the window
Min	Minimum value in the window
SD	Standard deviation of values in the window
Var	Variance of values in the window
Q1	Median of the lower 25 % of the values in the window
Q3	Median of the upper 25 % of the values in the window
IQR	Difference between Q3 and Q1

2) Feature Selection: To avoid over fitting, two feature selection models were used to find which features provide the highest prediction accuracy. Correlation Based Feature Subset (CfsSubset)[13] and Information Gain Attribute Evaluator[18] were used to find the best features, and based on this we concluded that "Mean-HR" (Mean Heart Rate) was the best feature. The five best features from the two evaluators can be shown in Figures (6) and (7). These features were obtained using Weka.

```
Attribute Subset Evaluator (supervised, Class (nominal): 10 Labels_Mod):
        CFS Subset Evaluator
        Including locally predictive attributes

Selected attributes: 1,5,6,7,8 : 5
                     Mean_hr
                     MAX_hr
                     MIN_hr
                     RANGE_hr
                     Median_hr
```

Fig. 6. CFS Subset Feature Evaluator

3) Classification Approach: The evaluation method is seen in Fig (8). There are two classification models used. The first one being intended-stress model, which was used for predicting intended-stress outcome variable based on the labels in Table (1). However, since different people perceive stressful activities differently, a self-reported stress model is also used where the mean stress level of all activities are calculated and then used to determine which activities were stressful and which

```
Attribute Evaluator (supervised, Class (nominal): 10 Labels_Mod):
        Information Gain Ranking Filter

Ranked attributes:
 1.1913   6 MIN_hr
 1.0432   5 MAX_hr
 0.9387   8 Median_hr
 0.8532   1 Mean_hr
 0.2939   9 Std_Deviation_hr
 0.2939   2 Variance_hr
 0.263    7 RANGE_hr
 0.0631   4 MCR_hr
 0        3 ZCR_hr

Selected attributes: 6,5,8,1,9,2,7,4,3 : 9
```

Fig. 7. Information Gain Attribute Evaluator

weren't.

Different Classifiers such as Random Forest[14], Naive Bayes[15], Logistic Regression[16] and Support Vector Machines[17] were used. And each time, different features were used. Mean-HR, Max-HR and Min-HR were used on each of the classification algorithms and each time the results varied. The algorithms were tested with 10-Fold Cross validation.

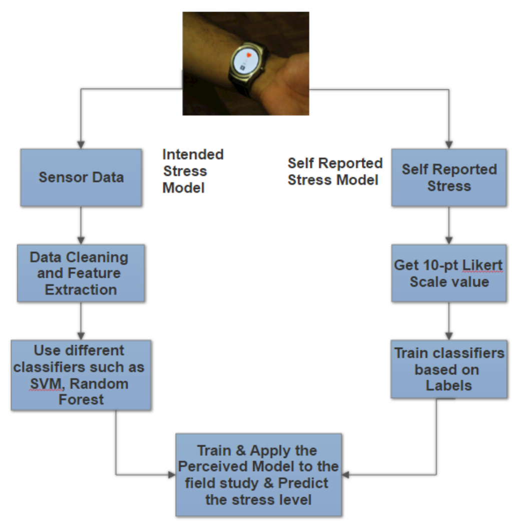

Fig. 8. Evaluation method

V. EXPERIMENTAL SETUP

The experiments were carried out in a lab setting where the participant was given very little information about the tasks to be performed beforehand. The participant was asked to wear the sensor and follow the instructions mentioned in the presentation. Each participant wore the smart-watch on the non dominant arm and performed the given tasks.

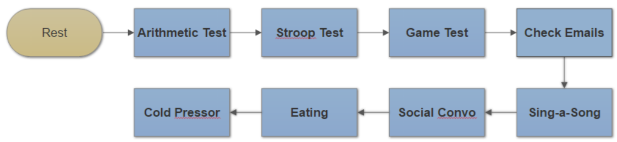

Fig. 9. Protocol followed by participants in the experiment

The protocol followed by the participants can be seen in Fig (9). Each block in the protocol represents a four-minute period and in between each of them, there was a two-minute rest period during which the participant was allowed to return to a normal resting state.

VI. Results

This section of the paper proceeds to discuss the predictive features along with the performance of the stress models

1) Classifier Accuracies: Four classifiers were tested to determine which provided the most accurate prediction. Naive-Bayes, Support Vector Machines, Logistic Regression and Random Forest were used on different features. For the intended stress model, the accuracy of Mean-HR can be seen in Table (3), and the accuracy of Max-HR in Table (4), and accuracy of Min-HR in Table (5).

TABLE III
RESULT BASED ON MEAN-HR PARAMETER

Classifier	Parameter	Weighted Avg
Random Forest	Mean-HR	74.3 %
SVM/SMO	Mean-HR	68.2 %
Logistic Regression	Mean-HR	67.9 %
Naive Bayes	Mean-HR	67.7 %

TABLE IV
RESULT BASED ON MAX-HR PARAMETER

Classifier	Parameter	Weighted Avg
Random Forest	Max-HR	90.1 %
Naive Bayes	Max-HR	66.9 %
Logistic Regression	Max-HR	64.1 %
SVM/SMO	Max-HR	63.4 %

TABLE V
RESULT BASED ON MIN-HR PARAMETER

Classifier	Parameter	Weighted Avg
Random Forest	Min-HR	90.0 %
SVM/SMO	Min-HR	67.8 %
Logistic Regression	Min-HR	67.6 %
Naive Bayes	Min-HR	67.1 %

Based on this data, we can see that Random Forest is the best Classifier and the most accurate feature is the Max-HR.

2) Self Reported Stress: Since different participants respond to stress differently, after each activity we asked each participant to rate the activity on a scale of 1-10. The tabulation of those values can be seen in Tables (6) and (7). And the average self-reported stress for each activity can be seen in Fig (9).

TABLE VI
SELF REPORTED STRESS TABULATION (PART I)

	CldPres	Email	Sing	Eat
Subject 1	6	4	10	1
Subject 2	8	3	9	3
Subject 3	7	3	8	1
Subject 4	6	3	10	2
Subject 5	6	3	9	1
Subject 6	3	4	7	1
Subject 7	8	4	10	2
Subject 8	7	3	6	1
Subject 9	4	1	9	1
Subject 10	2	9	1	2

TABLE VII
SELF REPORTED STRESS TABULATION (PART II)

	Math	Stroop	SocCon	Game
Subject 1	5	7	2	3
Subject 3	7	6	4	5
Subject 1	5	6	2	4
Subject 2	6	5	1	4
Subject 1	7	5	1	6
Subject 1	8	4	2	4
Subject 2	6	2	6	2
Subject 1	2	2	1	1
Subject 1	6	5	1	6
Subject 2	6	5	2	5

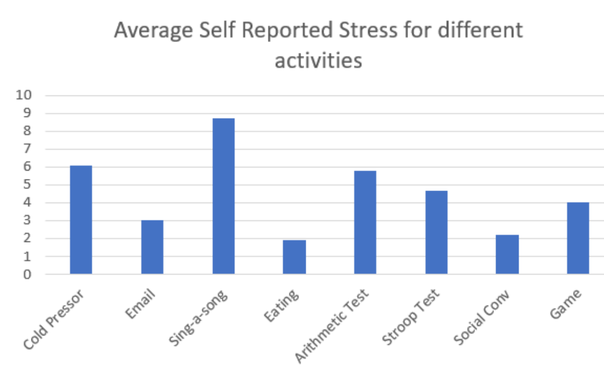

Fig. 10. Average Self Reported Stress values for different activities

Based on this data, we can conclude that "Sing-a-song" test is the most stressful activity.

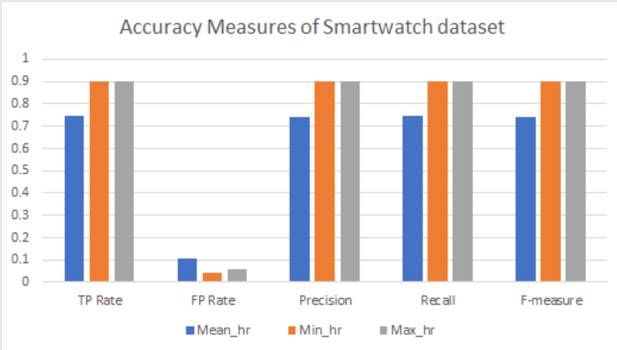

Fig. 11. Accuracy measures of Smart-watch dataset

TABLE VIII
CONFUSION MATRIX FOR MIN-HR

	S	NS	R
S	117	8	1
NS	4	49	7
R	0	7	45

TABLE IX
CONFUSION MATRIX FOR MAX-HR

	S	NS	R
S	123	2	1
NS	8	46	6
R	3	6	73

TABLE X
CONFUSION MATRIX FOR MEAN-HR

	S	NS	R
S	122	2	2
NS	4	27	29
R	3	28	51

The confusion matrix which shows the performance of the classification model for the different features extracted can be seen in the following tables. Table (8) shows the classification matrix for Min-HR, Table (9) for Max-HR and Table (10) for Mean-HR. The Accuracy measures of the Smart-watch dataset can be seen in Fig (11). The Recall for Max-HR is 90.3 %. The F-Measure for Max-HR is 90.3 %. The Precision for Min-HR is 90.2 %. The False Positive rate for Mean-HR is 10.6 %. The True Positive rate for Max-HR is 90.3 %.

VII. DISCUSSION

A. Training Phase

While performing the experiments, it was observed that the heart-rate of the participants was higher when performing a Stressful activity (S) when compared to a Non-Stressful Activity (NS) or a Rest Period (R). This observation was true for all the participants and for different activities. However, different participants showed different increases in heart-rate values for different activities. Eg: Some participants showed a greater increase in arithmetic tasks, while most showed a greater increase in Sing-a-song test. Another observation was that the heart-rate values from the smart-watch fluctuated if the watch was not worn properly.

B. Testing Phase

Once the data was collected from different participants and the model was built, it was time to feed these parameters on to the application to detect stress levels in real-time. The application was only tested in cases where the participant was not doing anything physically exhausting. He/she was either at rest, or doing some activity which was labeled as Stressful. Under all these constraints, the application was able to detect satisfactorily, whether a person was stressed or not and displayed the appropriate message on screen.

VIII. LIMITATIONS/FUTURE WORK

This work has few limitations and significant potential for future work. The accelerometers and the gyroscope of the smart-watch can be used in order to help detect the movements of a person and determine what type of activity a person is doing. This would help in further increasing the accuracy of the system. Also, the model has to be tested with a larger group of subjects to validate its accuracy.

Future work includes utilizing a Bio-Stamp Sensor, which is a commercial sensor. The data obtained from the Bio-stamp and the smart watch has to be correlated to test the reliability of the sensors and also to improve the accuracy of the overall system. The primary purpose of the Bio-stamp sensor is to detect type and degree of prenatal stress and also to analyze its effects on the baby. Recent studies [19] have shown that if a mother is stressed, anxious or depressed while pregnant, her child is at increased risk for having a range of problems, including emotional problems, conduct disorder and impaired cognitive development. Not all children are affected, and those that are, are affected in different ways. A proper methodology has to be implemented to identify stress and provide appropriate intervention.

IX. CONCLUSION

Different stress inducing experiments were performed on ten university students using a wrist-worn sensor. For a minute based classification, an accuracy of 90.12 % was achieved. The experiments show that Random Forest is the best classifier. There is sufficient proof to show that different in-lab experiments can induce stress with singing being the most stressful activity. The application was tested in a real-time setting based on the model and was able to satisfactorily determine whether a person was stressed or not.

ACKNOWLEDGMENT

We would like to extend our warmest gratitude to Prof Nabil Alshurafa and Northwestern EECS for giving us this opportunity to work on a such an interesting project. We are also grateful for the contributions of Dr. Michael Bass, Dr. John Rogers, Ha-Uk Chung, Dr. Judith Moskowitz and Dr. Lauren Wakschlag

REFERENCES

[1] Hovsepian K, al'Absi M, Ertin E, et al. cStress: towards a gold standard for continuous stress assessment in the mobile environment[C]//Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing. ACM, 2015: 493-504.

[2] Plarre K, Raij A, Hossain S M, et al. Continuous inference of psychological stress from sensory measurements collected in the natural environment[C]//Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on. IEEE, 2011: 97-108.

[3] http://www.stress.org/stress-iskilling-you/

[4] http://www.stress.org/what-is-stress/

[5] Begum Egilmez, Emirhan Poyraz, Wenting Zhou, Gokhan Memik, Peter Dinda, Nabil Alshurafa. UStress: Understanding College Student Subjective Stress Using Wrist-Based Passive Sensing

[6] Kurt Plarre, , Andrew Raij, Syed Monowar Hossain, Amin Ahsan Ali, Motohiro Nakajima, Mustafa alAbsi, Emre Ertin, Thomas Kamarck, Santosh Kumar, Marcia Scott 4, Daniel Siewiorek, Asim Smailagic, Lorentz E. Wittmers, Jr. Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment

[7] http://www.npr.org/sections/health-shots/2014/07/07/327322187/stressed-out-americans-tell-us-about-stress-in-their-lives

[8] Lu, H., Frauendorfer, D., Rabbi, M., Mast, M.S., Chittaranjan, G.T., Campbell, A.T., Gatica-Perez, D., and Choudhury, T., 2012. StressSense:Detecting Stress in Unconstrained Acoustic Environments Using Smartphones. In Proceedings of ACM Conference on Ubiquitous Computing.

[9] Vrijkotte, T. G.,Van Doornen, L.J., and De Geus, E.J., 2000. Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. In Proceedings of Hypertension

[10] Silverthorn, D.U., and Michael, J., 2013. Cold stress and the Cold Pressor Test. In Proceedings of Advances in Physiology Education.

[11] Brouwer, A.M., and Hogervorst, M.A, 2014. A new paradigm to induce mental stress: the Sing-a-Song Stress Test (SSST). In Proceedings of Frontiers in Neuroscience.

[12] Renaud, P., and Blondin, J.P., 1997. The Stress of Stroop Performance: Physiological and Emotional Responses to Color-Word Interference, Task Pacing and Pacing Speed. In Proceedings of International Journal of Psychophysiology.

[13] http://www.cs.waikato.ac.nz/ mhall/thesis.pdf

[14] Hall, M.A., and Holmes, G., 2003. Benchmarking Attribute Selection Techniques for Discrete Class Data Mining. In Proceedings of IEEE Transactions on Knowledge and Data Engineering.

[15] Sebastian Raschka. 2014. Naive Bayes and Text Classification I - Introduction and Theory

[16] https://www.stat.cmu.edu/ cshalizi/uADA/12/lectures/ch12.pdf

[17] http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf

[18] http://weka.sourceforge.net/doc.dev/weka/attributeSelection/InfoGainAttributeEval.html

[19] http://www.child-encyclopedia.com/stress-and-pregnancy-prenatal-and-perinatal/according-experts/effects-prenatal-stress-child