# Applying computational psychometrics to improve the reliability of psychometric tests

# Abstract:

*Recruitment tests predominantly consist of personality and aptitude tests. We probe into some of the key issues with these tests while looking for a solution to improve their reliability. Biosensors were found to be able to give cues to information about an individual emotional states and intellect. We propose introducing these sensors as an added criterion for determining the results of these tests. We intend to collect the data through biosensor and classify it with SVM and CNN following which we will conduct some statistical analysis to check its reliability.*

# Introduction and Problem Definition:

It is no secret that efficient handling of human resource plays a pivotal part in determining the future of an organization. This process will usually always begin with choosing the right candidate which is easier said than done considering the vast amount of factors that come into play. ([Hacker](#ref_1), 1997, p. 13) US department of labour states that poor hiring procedures can cost up to 30% of the employee’s first year’s wages. A poor hire maybe damaging to the work culture and reduce morale among employees. Recruitment specialist [Half](#ref_2) orchestrated a survey in New Zealand and found that about 72% of HR managers were not pleased with their workforce. He states that there was a productivity loss of close to 49% and a loss in morale of up to 30% resulting from a bad hire (2016). Recruitment procedures tend to vary a lot based on the size of the organization which could be owed to factors like the complexity of the organization, requirement of human resource and the lack of a department specific to human resource. Large firms generally have a well-defined human resource department as opposed to small firms which are more likely to let someone in a managerial position handle the recruitment procedure or simply hire the services of a professional recruitment agency. Small firms are also more likely to depend on referrals from existing employees. Emphasis is given more on the work culture of the firm and fitting in with the current staff. This leads to employees being promoted and given additional tasks which eventually results in the job including a collaboration of very specific tasks thus increasing the dependence on the said employee. The replacement of such an employee and the inevitable training period in such a case is more than likely to cross a drop in productivity, morale and monetary loss. An average recruitment process will go through some standard procedure like an analysis of the job related specifications, tenure, work-culture followed by job advertisement, CV shortlisting and the interview which may or may not involve some kind of assessment tool like psychometric tests or aptitude tests. ([Brady et al.](#brady), 2017) Most large scale and medium scale companies have a tendency to rely on some combination of psychometric, [cognitive](#cognitive) or aptitude and ability related tests based on the specifications of the job. Psychometric tests are devised to measure the personality of the candidate. This sort of testing enables the hiring team to check if the candidate’s personality matches with the job’s requirements and the company’s culture. Aptitude tests the candidate’s cognitive skills. It is based on the idea that the intellectual performance of an individual maybe divided into many subsections like numerical skills, verbal reasoning, [abstract reasoning](#asbtract_reasoning), [spatial reasoning](#spatial_reasoning) and [mechanical reasoning](#mechanical_reasoning). Candidates maybe dishonest with personality tests to fit the company’s notion of an ideal candidate. ([Donald](#ref_3),2010) There is also an increasing debate over the validity of psychometric tests due to its lack of supporting evidence for example the public service association in New Zealand deemed psychometric testing as pseudoscience and attempted to take legal action against Inland Revenue Department over its use of psychometric tests. A controlled study conducted by [Miller and Barrett](#ref_4) found that candidates receiving training for personality theory were likely to score higher on psychometric tests devised for safety force personnel and were also more likely to get the job (2008). A meta-analysis conducted by [Dakin](#ref_5) found personality tests to be inefficient on its own for managerial positions. They argue that a large part of this is due to the hiring team’s inability to select desirable qualities for the role in question. They argue that it could not be considered as a precursor to success even when the personality profile was well defined (1994). Higher scores on aptitude tests don’t necessarily indicate the ability of a candidate to succeed as there is a possibility of them practicing these tests and learning known patterns to obtain a higher score. There is also a possibility of weening off suitable candidates that were anxious during the test. The consequences of a poor hire when it comes to top level executives like the CEO can be especially dire.

# Biomedical Perspective:

## 2.1 Biosensors:

A biosensor is any biological sensor capable of obtaining and analysing biological data from organic elements and converting it to electronic signals. We are considering using heart rate variability(HRV), electro-dermal response(EDR) and skin temperature sensors for our research since they are cheaper and more efficient when compared to other sensors. HRV and skin temperatures provide consistent results due to its lower level of abstraction. The quality of the data is higher due to its implications on the sundry of bodily systems. Electroencephalography is capable of providing detailed information about the brain but is obfuscated behind complexity and noise in the data. Cortisol levels are usually measured by collecting bodily fluids and commercial sensors for obtaining real time data through interstitial fluid is not available of yet. Electromyography (EMG) sensors used to measure electrical activity can track stress levels to a reasonable extent but would be limiting for our research due to the discomfort and ambulatory difficulties caused by these sensors. HRV sensors measure the variance in the time period between heart beats. It is one of the most efficient means of measuring changes in emotional states, stress and cognitive load due to its effect on the nervous system and thus the brain. Electro-dermal sensors measure the electric properties of the skin and have been historically used in polygraph test. It is considered to be efficient due to its effect on the sympathetic nervous system. However, it has been a frequent subject of scepticism among researchers due to its inconsistency. This can be corrected by observing the EDR response over the course of a few days or less and establishing a threshold. Skin temperature has been linked to increase in stress and cognitive load. These findings have been discussed further with sources below in their corresponding subsections.

### 2.1.1 Heart Rate variability (HRV):

Heart rate variability or HRV is a measure of the variance of time between consecutive heart beats. This variance between beats taking spikes into consideration is measured in milliseconds and is called as the RR interval, while NN (normal to normal) interval can be defined as the same with emphasis on normal beats instead of spikes. It is not to be confused with heart rate (HR) which is the measure of the number of contractions within a given time period. The abnormal spikes in Figure 1 describe an RR interval.

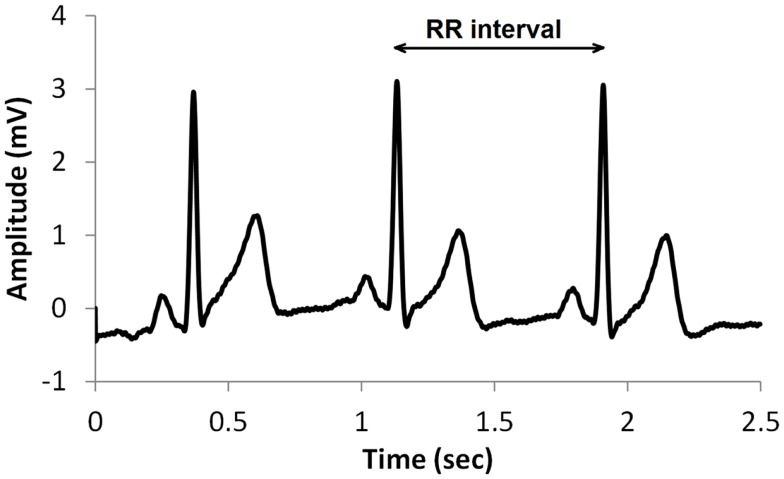


Figure . RR Interval ([Cornforth](#ref_6) et al, 2014) The large spikes in this figure are called R-waves. RR intervals can be described as the time difference between two consequent R-waves.

HRV is usually measured by an ECG (Electrocardiography) device as it is considered to be more accurate, however, a lot of portable devices are also available in the market which could do the same. HRV as opposed to HR is capable of measuring physiological functions besides [cardiovascular](#cardio) activity. It might also be used as means to measure [autonomous nervous system (ANS)](#ANS) and Respiratory system activities. ANS activity is evaluated by observing [sympathetic](#sympathetic) and [parasympathetic nervous system](#parasympathetic) activities which are again dependent on HRV. Some of the parameters involved in analysing HRV based on time and frequency domains are as stated below

* SDANN or standard deviation of 5-minute average NN intervals of entire recording and indicates short term variability
* SDNN or standard deviation of NN includes both short term and long-term variability.
* rMSDD is the squared root of mean squared differences of consecutive NN intervals. It indicates parasympathetic nervous system activity.
* meanNN indicates the mean value of NN interval.
* LF which is the low frequency component lies between 0.04-0.15 Hz and is equated with both sympathetic and parasympathetic nervous system activity.
* HF is the high frequency component and lies between 0.15–0.4 Hz. It represents parasympathetic nervous system, respiratory system and blood pressure.
* LF/HF ratio is the ratio between LF and HF frequencies.
* VLF/P stands for very low frequency power and lies below 0.04 Hz.
* P stands for total power of the frequency spectrum.
* RSA or Respiratory sinus arrhythmia is the change in HRV corresponding to respiratory cycle.

HRV is generally considered as the best biomarker of psychological stress besides [cortisol](#_Cortisol:) levels. Increase in sympathetic activity causes a decrease in the time between heart beats and, on the other hand, an increase in parasympathetic activity is co-related with the increase in the time interval between the successive beats. High frequency HRV shows a parasympathetic influence while Low frequency HRV is noted to be connected with sympathetic nervous system activity. The amygdala is responsible for the flight or fight response and people with anxiety disorders, Post Traumatic Stress Disorder (PTSD) and phobias show a higher [amygdala](#amygdala) response. In ([Thayer et al](#ref_7), 2012) It has been suggested that hyper activation of the amygdala is co-related to lower HRV among other factors. [Appelhans and Luecken](#ref_8) observed the connection between emotional regulation and HRV (2006). They have noted that an increase in emotional regulation is connected to an increase in resting HRV. A study conducted by [Hjortskov](#ref_9) et al reports that cognitive and emotional stress correspond to a decrease in HRV. Participants were asked to perform certain tasks on a computer and three different sessions were conducted. In the introductory and stress sessions the host was strict towards the participants and behaved in an unfriendly demeanour towards their colleagues. These factors of stress were removed in the control session. HRV was measured throughout the sessions with ECG. Blood pressure was also monitored. The participants showed an increase in parasympathetic activity during the control session and an increase in sympathetic activity during the stress session. They found that blood pressure did not seem to have a strong influence in both sessions. This experiment indicates that psychosocial stressors have an influence of HRV (2004). [Thayer et al](#ref_10) have hypothesized that the [Vagus](#vagus) nerve responsible for parasympathetic activity has an influence in the functioning of the prefrontal cortex which is responsible for some important cognitive functions like decision making, social behaviour and planning and thus HRV can be considered to have an important role in modulating cognitive behaviour. They have given evidence of this via neuroimaging. An experiment involving measurement of [cerebral](#cereberal) blood flow and emotional modulation was conducted and the participants seemed to show a decreased activity in the prefrontal cortex and a lower HRV was observed thus suggesting that cognitive functioning is reduced during emotional distress. They have also managed to co-relate attention span, [working memory](#work_mem) and [situational awareness](#situational_aware) in a different set of experiments (2009).

### 2.1.2 Electro-dermal Response (EDR):

Electro-dermal activity (EDR) or galvanic skin response (GSR) maybe defined as the change in the electrical properties of the skin. These properties include electric potential, resistance and conductance of the skin. The generation of EDR can be attributed to the sudoriferous glands responsible for producing sweat. It can be measured with an ECG or [Electromyography (EMG)](#emg) device by measuring the variance in the electrical conductivity or resistance of the skin. GSR is influenced by the autonomic nervous system just like HRV as sweating is controlled by the sympathetic nervous system. An increase in sympathetic activity causes an increase in skin conductance. EDR is considered to be responsive to emotional arousal due to this. ([Iacono](#ref_11), 2001) It has been previously used for polygraph tests and has been deemed ineffective as might be easily faked by various means like clenching your fists. It can be variable individually which makes it hard to use as a biomarker for stress. However, with the advent of [machine learning](#machine_learn) we could consider the possibility of EDR being viable to a reasonable extent. A study conducted by [Shi et al](#ref_12) linked EDR to cognitive overload. The experiment showed comparisons between unimodal and multimodal [human computer interfaces (HCI)](#hci) using gesture and speech input. The mean EDR levels appeared to drop after the subjects gained a level of expertise with the interface. EDR was found to be proportional to the increase in cognitive overload and multimodal HCI showed the least amount of mean EDR (2007). ([Villarejo, Zapirain, Zorrilla](#ref_13), 2012) A stress detection sensor was developed based on EDR and controlled by ZigBee which is an IEEE 802.15.4 based wireless technology that is commonly used for Internet of things (IOT) projects due to its low power consumption. The study considered the individual variability of EDR and devised a threshold for each subject participating in the study. The subjects were exposed to different levels of stress with simple exercises like performing mathematical operations, fast reading and staying still. The tests were verified with three different machine learning algorithms i.e. [Bayesian Network](#bayesian_networks), [Decision Trees](#decision_trees) and [Sequential Minimal Optimization (SMO)](#smo). The subjects seemed to show different levels of EDRs to each exercise and the lowest level of EDR was observed when the subject stayed still, thus connecting EDR to their level of stress. In ([Kurniawan, Maslov, Pechenizkiy](#ref_14), 2013) an experiment was devised to classify stress levels by combining EDR with speech output using machine learning algorithms. The data was obtained by subjecting the participants to [Stroop-Word colour congruent](#stroop) and arithmetic tests. The machine learning classifiers included [Gaussian Mixture Model (GMM)](#gaussian_mix_model), [Support Vector Machines (SVM)](#svm), [K-means classification](#kmeanscluster) and Decision Trees. SVM showed the highest level of accuracy while K-means exhibited the lowest amount of accuracy. They observed that the results varied individually and differed on a day to day basis. They assumed that the reliability of their system is due to the inconsistencies in the EDR data. [Collet et al](#collet) orchestrated a study where they tried to understand the influence of autonomic nervous system to basic emotions. The ANS parameters consisted of skin temperature, skin conductance, skin resistance and skin blood flow. skin resistance and skin conductance had an inverse effect on each other. They report being able to measure basic emotions while also being able to make a clear distinction between fear and sadness, surprise from disgust and being able to classify them according to a negative or positive with skin resistance (1997).

### 2.1.3 Skin Temperature:

The human skin consists of three different layers and is responsible for certain homeostasis functions like maintaining skin temperatures, insulation and water balance. Variance in skin temperature is medically perceived as a symptom of disease. However, skin temperature can deviate to a certain extent from external factors like temperature and internal factors like psychological distress or intake of food and drink that is high in calories. These changes in temperature are generally short term. Skin temperature is likely to drop with acute stress from [vasoconstriction](#vasoconst). [Lisetti and Nasoz](#ref_15) tried to determine human emotion through the use of a wearable sensor (BodyMedia SenseWear) and skin temperature along with EDR and heart rate were the factors involved in determining the emotional states. Subjects were shown some movies with the arm sensor attached to them. The data was normalized by averaging the data collected during the relaxation session. The data was classified using certain machine learning algorithms and the features included min, max, mean and variance of the [normalized](#normalization) data. [K-nearest algorithm (KNN)](#knn), [Discriminant function analysis (DFA)](#discrim_fun_an) and [Marquardt’s backpropagation algorithm (MBA)](#marq_back_prop) were applied. Their model obtained different accuracies for each emotion. The overall accuracies for KNN, MBA and DFA were 72.3%, 84.1% and 75% respectively. They were able to relate EDR with frustration (2004). Another study conducted by [Kaufman and Kagan](#ref_16) have managed to link emotional states with skin temperature. The skin temperature of thirty-two student’s hands were monitored while being subjected to personal questions to evoke negative emotions and upbeat films to evoke positive emotions. They found that the skin temperatures increased in the former case, while it dropped in the latter case (1996). [Khan and Ingelby](#ref_17) have tried to understand the effect of human emotions on thermal features of the face. They obtained the thermal facial image data through [infrared thermography](#inf_therm). The features were divided into 75 pieces based on the location. This data went through median smoothing filter following sobel operator based edge detection algorithm during the image processing procedure. Feature classification was done by means of [Linear Discriminant Analysis (LDA)](#LDA) and initially the accuracy was found to be variable based on the amount of facial expressions. It was found to be quite low at 57% with over six facial expressions and up to 83.3% with just facial three expressions (2009). [Puri et al](#ref_18) reported that human psychological stress levels can be connected to increase in blood of the frontal blood vessel of the forehead. An experiment was conducted wherein subjects were asked to take the Stroop colour word conflict test to measure their stress levels. The experiment was conducted in two sessions; the participant’s energy expenditure was calculated with a device in the first session followed by measurement of their forehead temperature through infrared thermography. The data from both sessions were normalized and the data obtained from the first session was used for validating the change in stress levels. The difference in the values obtained from both sessions was found to be no more than eight units which they consider to be reasonable enough to be a viable solution for computing stress levels (2005). In ([Pavlidis, Levine, Baukol](#ref_19), 2000) it has been found that variances in facial temperatures between calm and anxious states of mind are drastic. The thermal patterns in individuals was found to be distinct with the onset of anxiety. Anxiety appeared to induce periorbital warming with cooling of the cheek area. This can be owed to the increase in blood flow to the eyes from the cheek. [Allison and Schnell](#ref_20) were able to establish an affiliation between facial temperatures and cognitive workload. They measured the changes in temperatures of eyes, chin and cheek area in response to three different levels of [cognitive workload](#cogn_load) in their test subjects and observed the changes to be minute in low and medium levels of cognitive workload however they were able to distinguish the changes accurately after passing the features through an [Artificial Neural Network (ANN)](#ANN) classifier with a hyperbolic tangent function as a hidden layer with an accuracy above 97% when data from a single test subject was considered. The accuracy decreased and was found to be within the range of 76% to 86% in the case where data of all the test subjects were included which they reckon is due to the participants having different cognitive thresholds (2010).

### 2.1.4 Cortisol:

Cortisol is a [steroid](#steroid) hormone which aids in [gluconeogenesis](#gluco) to increase blood sugar, maintain fluid balance and blood pressure in response to a stressful event. It is one of the three stress hormones besides [adrenaline](#adrenaline) and [norepinephrine](#noradrenaline) that is released by the adrenal gland. Cortisol levels are considered to be the best biomarker for stress. Cortisol levels in human can be easily monitored by measuring cortisol levels in hair strands or saliva. One of the major drawbacks to these methods is the inability to obtain continuous data in response to stress. This has led to the search for a better approach to quantify real-time stress response while still trying to keep the quality of data that cortisol levels have been known to offer through non-invasive means. [Interstitial fluid (ISF)](#interst_flu) can be found in between tissue cells and constitute 16% of the body weight. It consists of neurotransmitters, proteins, sugars, salts and hormones including cortisol. The cortisol levels found in ISF is said to be higher than salivary levels of cortisol. [Venugopal et al](#ref_21) have developed a sensor to continuously monitor cortisol levels via the interstitial fluid by applying vacuum pressure on the micro pores of the outermost layer of the skin coupled with an electrochemical cortisol detection system. The procedure was reported to be free of pain and the cortisol levels were found to be about four times higher than salivary levels (2011).

### 2.1.5 Electroencephalography (EEG):

Electroencephalography (EEG) is used to measure the electrical activity in the brain using electrodes placed right above the scalp. The electrical activity is the result of Ionic charge within neurons. It has been used for measuring [epilepsy](#epliep), migraines and brain activity within coma patients. EEGs are non-invasive in nature compared to traditional [brain computer interfaces (BCI)](#bci) and can provide information on the electrical activity within the brain which has resulted in extensive use for neuroscience research. EEGs could provide an insight into an individual’s emotional states and cognition. [Berka et al](#ref_22) reported that the rise in EEG workload can be correlated to the rise in mental workload. Subjects were asked to perform five different forms of mental tasks and EEG activity was recorded. The EEG workload was found to rise in a linear fashion to arithmetic and memory tasks. The measures were found to plummet during tasks that required vigilance as the time progressed even though the workload was kept constant hinting that attention may also be measured with EEG (2007). Emotion recognition with EEG is a challenging task and still considered to be in its infancy. In ([Lin et al](#ref_23), 2010) an experiment was conducted in an endeavour to classify four basic human emotions (Joy, Anger, Sadness, Pleasure). Subjects were asked to listen to emotion evoking music while their EEG data was recorded. The features were extracted with [Short Time Fourier Transform (STFT)](#stft) and classification was attempted with Support Vector Machines (SVM) and [Multilayer Perceptron (MLP)](#multilay_percep) and the highest accuracy was observed for SVM at 80.86%. [Hou et al](#ref_24) made an attempt at monitoring psychological stress with EEG. Participants were required to take the Stroop colour word test and the data was classified at 4 levels of stress with Support Vector Machines (SVM) and K-nearest neighbour’s algorithm (KNN) and they achieved an accuracy of 85.15% and 76.72% respectively (2015).

# Human Resource Perspective:

## Computational Psychometrics:

Computational Psychometrics is an emerging field that incorporates computational sciences with psychometrics. It has shown potential in the fields of clinical psychology, rehabilitation and artificial intelligence. Machine learning and mathematical models could simulate human behaviours on different levels and the information could be useful in disciplines concerning psychology. Integrating Biosensor data with computational psychometrics would increase its potential by providing a self-checking feedback mechanism. [Cipresso](#ref_25) has applied computational psychometrics with virtual reality technology to study human behaviour in situations that are hard to replicate in real life. He argued that virtual reality provides a strong sense of presence of self and can be beneficial to understand human behavioural dynamics. He created a virtual environment with NeuroVirtual3D, which is a tool often used for [Virtual Reality (VR)](#VR) based research due to its simplicity. Subjects could interact with the virtual environment through Kinect, which is a real-time tracking device developed by Microsoft for its Xbox gaming consoles. The participants were given different kinds of sensory stimulus and cognitive tasks and a hypothetical condition was introduced to monitor their response to stress. Biosensor data was obtained through EEG, EDR, EMG (electromyography) HRV and respiration sensors and synchronized with MATLAB algorithms. He obtained a stress threshold for each participant and observed that women cope better with academic stress (2015). [Villani et al](#ref_26) studied the impact of communication on decision making in a hypothesized scenario testing the ability of the subjects to trust each other. They classified the features based on emotional valence and arousal models. They obtained the biosensor data by means of EEG, EDR, [Blood Volume Pulse](#BVP) (BVP) for HR and HRV, eye tracking, EMG and respiration sensors. Participants were placed in different rooms and made to play an investment game with monetary incentives that tested their trust in each other while communicating through a video call. They connected the increase in arousal with monetary gain and the ability to stay relaxed could not be linked to trust in the experiment (2015). [Segkouli et al](#ref_27) developed a computational simulation to understand changes in cognition, perception and behaviour arising from dementia (2015). [Polyak, Davier and Peterschmidt](#ref_28) developed a game to measure collaborative problem-solving skills. The game involved a series of puzzles and the participants were allowed to communicate over a chat box. The data obtained from the chat was logged along with means for identification and time. Bayesian evidence training algorithm, KNN and clustering algorithms were used for data classification and their collaborative problem-solving skills were scored (2017).

## 3.2 Five Factor Model:

Human personality has been a subject of interest for centuries. Researchers have pondered on how to classify and measure personality traits without reaching a conclusion. The last few decades, however, had seen a major development in personality research and we no longer suffer from the paucity and unreliability of empirical data([McCrae and Costa](#costa),1988; [Digman](#digman), 1990). The five-factor model which is commonly accepted as the best descriptor arose as a direct result of this. The five factors include

* Openness to experience: It refers to the predisposition of a person to seek new experiences. It is characterized by the intellectual curiosity and ability to appreciate new perceptions. ([McCrae](#mccrae),1987; [Kausch et al.](#kausch),2006) Individuals with a higher score on this trait tend to be self-reflective and creative although prone to risk-taking behaviours. These individuals could vary from being extremely paranoid to being gullible. A low score on this trait may imply narrow mindedness, scepticism and pragmatic behaviour.
* Conscientiousness: It is characterized by the ability of a person to exercise self-restrain and work towards their goal in an organized fashion. ([Smits et al.](#smits), 1993; [Busato et al.](#busato), 2000) Conscientiousness generally implies a higher degree of academic and work-related success. It is a trait commonly found among workaholic people.
* Extraversion: It is a trait found in individuals with a higher need for socialization and expression. Extraverted people may show high skill at maintaining interpersonal relationships but may also come out as domineering.
* Agreeableness: This trait is characterized by an individual’s predisposition with virtues like compassion, empathy and compliance. ([Graziano et al.](#graziano), 1996) People with a higher score on this trait may come out as passive with a tendency to avoid conflict but it also associated with a higher capability of conflict resolution.
* Neuroticism: It can be tethered to the inclination of an individual to stay in a negative state of mind. People who score high on neuroticism show higher levels of anxiety, depression, insecurity and emotional instability.

It has since been used to study the role of personality and its influence. [Rothmann and Coetzer](#rothman) tried to establish the impact that the big five personality traits had on job performance. The survey design was cross sectional and consisted of a mix of pharmacists and non-pharmacists selected from over 30 pharmacies. They administered NEO-PI-R tests and found that high scores on extraversion, openness to experience, conscientiousness traits and low scores on neuroticism traits can be linked to a performance increase of about 15%. They suggested that the influence of these three personality traits are more significant on management performance (2003). [Giberson et al.](#giberson) found that CEO’s personality traits had an influence on the cultural values of the company after conducting a study with over 33 organizations. They obtained relevant information regarding the company and the employees were asked to complete an anonymous survey with questions regarding their CEO based on the five-factor model. The company values were obtained with the [CVI](#CVI) (Competing values instrument) scale. Relationships were observed after comparing the data obtained using ([Cheung and Keeves](#cheung), 1990)[HLM](#HL) (Hierarchical Linear Modelling) They found that the CEO’s emotional stability (low neuroticism) and agreeableness played the most significant role in shaping the company’s culture. Agreeableness was found to positively effect on how the employees viewed their company culture being marked by clan values while at the same time it had contradicting effect on how they viewed adhocracy and market values. A CEO with higher agreeableness indicated stability along with a higher staff morale while conversely a CEO with lower agreeableness results in a company culture that is performance based (2009). [Nadkarni and Herrmann](#nadkarni) suggested that an organization’s performance is directly dependent on strategic flexibility which in turn is dependent on CEO personality. Strategic flexibility can be defined as the ability of a firm to adapt to new environments. They backed their claims with evidence they obtained after conducting a macro study with over 217 firms. The data was obtained through the means of personality surveys filled by the CEOs which were based on the five-factor model and some strategic flexibility surveys. They established some control variables after considering the size of the organizations and firm performance was measured by monitoring the company’s financial records. The data was modelled with different statistical methods and the relations were observed. They observed that each trait in the five-factor model had an influence on the strategic flexibility of the firm which in turn affected firm performance. Conscientiousness had the most diminishing effect on the performance of the company while low neuroticism, openness to experience and extraversion ameliorated the firm’s performance (2010). [Booth et al.](#booth) conducted a study where they tried to measure the personality facets of CEO and tried to understand the difference between CEO and managers at different levels. They assumed that the differences would be significant owing to the vast differences in recruitment procedures. The data was acquired through a development centre with over 3000 participants. They were split into three categories based on the roles they were currently playing at work. The categories consist of non-managerial, managers and top-level executives, which was again divided into CEO and non-CEO. The participants were then evaluated using the NEO-PI-R, which is a personality inventory based on the five-factor model consisting of 240 items. The NEO-PI-R personality inventory divides the five factors into over 30 distinctive facets. Participants were also requested to take the graduate managerial assessment test. They observed that CEOs showed lower level of impulsiveness and vulnerability while showing a higher level of dutifulness, competence, self-discipline and activity. They observed that CEOs did not possess different traits when compared with other managers and instead the only seemed to score higher on these traits in comparison (2015). Transformational leadership is a style of leadership where the leaders intend to motivate their subordinates by serving as a positive role model. It assumes that there are four elements that make a good leader which include idealized influence, idealized motivation, idealized stimulation and idealized consideration. [Judge and Bono](#judge) orchestrated a study where they tried to understand the influence of personality on transformational leadership. About 539 people participated in the survey of which 316 where current students taking a community leadership program while the rest of the 223 participants included Alumni from the same institution. Participants were asked to take a Multifactor Leadership Questionnaire and the NEO-PI-R which is based on the five-factor model. Additionally, the effect that leaders had on their subordinates were observed through the means of surveys catered to measure the effect that the leaders had on them in accordance with the four dimensions of transformational leadership. To their surprise, they discovered that Agreeableness had the highest influence on transformational leadership. Extraversion and openness to experience had a secondary influence on transformational leadership while neuroticism and conscientiousness had almost no effect (2000). [Palaiou and Furnham](#palio) tried to uncover the differences between CEO and staff consisting of five different sectors which included engineering, finance, legal, human resource and marketing sectors. The study was conducted on a macro scale with a total of 16,258 participants out of which 7,768 were CEOs. The tests administered were based on the NEO-PI-R used to measure the big five personality traits. Their findings established that CEOs scored low on neuroticism and highest on assertiveness and industriousness facet belonging to conscientiousness trait compared to professionals from all the other five sectors. CEOs also showed lower enthusiasm than marketing professionals although it was found to be higher than engineers. CEOs also showed lowest openness to experience compared to the rest when it was personal. The differences in CEO’s personality facets were found to be the highest when compared to engineers and lawyers (2014). The NEO-PI-R personality inventory has been frequently used to devise recruitment tests due to its empirical resilience. ([Rosse et al.](#rosse), 1998) However, the issue of response distortion persists. They established this in their study which comprised of 197 job applicants and 73 job incumbents. The participants were asked to take a take a modified version of NEO-PI-R which combined the BIDR-IM (Balance inventory of desirable responding) which is commonly used to log deception on psychiatric tests. They found that response distortion was higher among job applicants after comparison with job incumbents. The distortion was also found to be closely related to the ideal personality facets of the job they were applying for. The applicants were found to have significantly higher scores compared to the job incumbents. [Watson and Clark](#watson) suggested that emotional valence has a direct correlation with the personality traits observed in the five-factor model. They corroborated this claim in their study where they obtained personality data with the NEO-PI-R and compared it with the emotional influence. The participants consisted of university students and the emotional affect was classified with the [PANAS](#panas) (Positive and Negative affect schedule) which was filled alongside with the NEO-PI-R. They found that positive affects had a strong influence on extraversion and conscientiousness. openness to experience and agreeableness were also found to have a stronger correlation with the same, however, they also showed some correlation with negative influence. Neuroticism was found to have a strong relation with negative emotionality. These results were mostly one dimensional indicating that they are independent in nature (1992). [Judge et al.](#judge_leadership) also established a connection between negative emotionality and Neuroticism in their meta-analysis concerning the effect of leadership on personality traits. They also found a direct correlation between openness to experience, agreeableness, conscientiousness and extraversion (2002). These results were once again recreated in an organizational commitment study conducted by [Panaccio et al](#panaccio). Additionally, they were also able to confirm that agreeableness was positively related to positive emotionality and negatively with negative affect (2012).

# Machine Learning:

Machine learning is a field of computer science that is concerned with devising and applying algorithms that enable computers to learn without being explicitly programmed to do so. It allows computers to learn through experience obtained by analysing data and use it to make predictions. Machine learning algorithms can be divided into three main categories

* Supervised machine learning – Supervised machine learning algorithms go through a training and test period. In this type of machine learning, we provide the classifier with an input data set and labels along with the actual output during the training period. The classifier learns the patterns which can be applied during the test phase to predict the output which it is not supplemented with. One may compare the predictions obtained from the training set against the actual output to find the accuracy of the classifier. Supervised machine learning can be further classified into two groups.
  + Regression: A regression model consists of input and output data that are continuous. It is used in cases like housing price prediction where the output variables are a continuous numeric quantity rather than a fixed class.

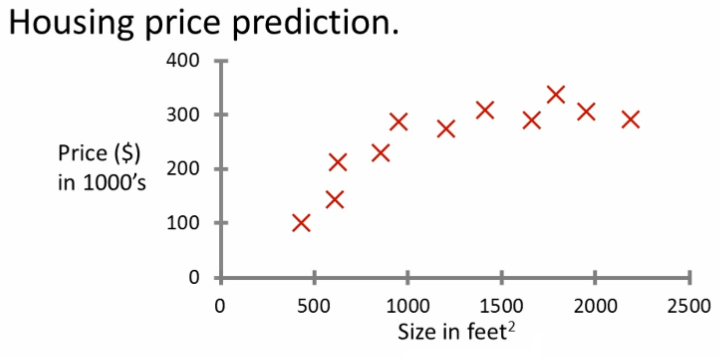


Figure 2 Regression([Holehouse](#holehouse), 2012) The image describes a regression problem wherein the input data on x and y axis are continuous. X marks the output variables and the values are continuous decimal numbers.

* + Classification: A supervised learning problem is categorized as a classification problem when the data is to be classified into discrete categories. The classification problems are often binary in nature i.e. classifying mail as spam and not spam although it can also be multivariate.

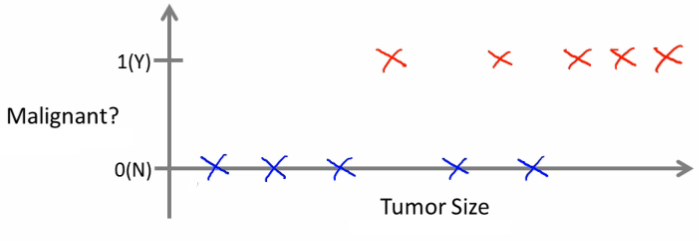


Figure 3 Classification([Holehouse](#holehouse), 2012) The image describes a binary classification problem wherein we intend to identify if a tumour is benign or malignant based on its size. The red mark categorizes it is as malignant.

* Unsupervised Machine learning: It comes with no labels for the response variables. There is no need to train the classifier as it clusters the data and finds patterns on its own. The entire process is unsupervised, this is useful while working with new data wherein no inherent pattern has been identified.
* Reinforcement learning: It is used in cases where there isn’t a definite solution. It utilizes a reward system which is activated when the classifier performs well. Reinforcement learning has applications in control theory and video game development.

There have been various attempts at classifying human emotions hinged on the concept of emotional valence and arousal with the use of machine learning. A huge portion of this research comes from studies concerning HCI (Human Computer Interaction). [Mandryk and Atkins](#mandryk) discovered that the fuzzy systems achieved an accuracy at classifying emotions that was very similar to that of manual classification in their HCI study. The features included heart rate, GSR and EMG data obtained while their participants played video games. The [fuzzy systems](#fuzzy) were developed on the valence and arousal model. EMG was used to detect valence from facial expression. However, they deemed it as weak and included HR as an additional control for valence. Arousal was monitored exclusively through GSR and HR. Fuzzy systems provided an accuracy of around 65% for both valence and emotion while a manual approach had an accuracy of 65% for arousal and 47% for valence thus suggesting the usage of mathematical and machine learning techniques over a manual approach (2007). [Yannakakis and Hallam](#yannakakis) designed an entertainment model using preference learning to determine what kind of video games children classify as fun. The monitored HRV, HR and skin conductance to create an artificial neural network which classified the data with an accuracy of 66%. They observed that high frequency band obtained from BVP was the strongest biomarker of entertainment (2007). [Chanel et al](#chanel) framed an interactive game that adjusted the difficulty level based on the user’s emotions. This was done by classifying HR, GSR, blood pressure, respiration and temperature by SVM with a radial function in accordance to the valence and arousal model. They achieved an accuracy of 53.33% at classifying the data into the categories of positive engagement, boredom and anxiety (2008). [Lee et al](#lee_yoo) propose the usage of neural networks for classifying emotions after they achieved an accuracy of 80.2%. They measured HRV and GSR data from their participants as they watched video clips. The emotions were classified into four dimensions which were sad, calm, pleasure and fear (2005). [Kim and Andre](#kimandre) contrived a similar experiment where they tried to classify emotions in response to music. The biosensors included EMG, GSR, ECG and respiration sensor. The data was classified utilizing LDA combined with [SBS](#sbs) (sequential backward selection) due to the complexity of their data surging from an excess of features. They achieved an accuracy of 95% and they found that a lot of the features were person specific (2008). [Ferdinando et al](#ferdinando) achieved a relatively lower accuracy for emotion classification when they tried to use HRV data obtained through ECG signals from an emotional response dataset. They used SVM and obtained an accuracy of 47.69% and 42.55% for arousal and valence, respectively (2014). However, [Kim et al](#kimkim) managed to obtain a higher accuracy when they try to classify emotions into four categories with SVM. They obtained an accuracy of 78.84% with three categories and 61.8% with four categories. Their dataset consisted of HRV, skin temperature and EDR which was obtained by asking their participants to gaze at a toy after telling them stories to evoke emotions (2004). [Oliveira et al](#oliveira) attained a similar accuracy for their classifier at 69% with five categories when they used SVM. The accuracy was far lower at 47% for k-NN. The biosensors used in this case included HRV, HR, GSR and respiration and the medium of evoking emotion included movie clips (2011). [Martinez et al](#martinez) used [CNN](#cnn) (convolutional neural networks) to classify emotional response to video games. The features consisted of blood volume pulse and skin conductance. The accuracy plummeted when they tried to use the features in an ad-hoc manner, however, it rose up to 75% when they tried to fuse it into one entity (2013). [Jang et al](#jang) tested the performance of different machine learning techniques at classifying simple human emotions. They used a wide array of features which were derived from ECG, GSR sensors. They also used temperature and BVP sensors. [Naïve Bayes](#naive) attained the highest accuracy at 53.9% followed by LDA at 52.7% and CART (classification and regression tree) at 42.7%. SVM had an accuracy of 39.2% which was the lowest among the four (2014). [Valenza et al](#valenza) conducted a study where they tried to classify affective emotions using the valence and arousal model. They only chose participants that were of sound health and they used images as a medium of emotion elicitation. They main purpose of the research was to understand what kind of features were better at classifying emotions. The features consisted of many elements derived from ECG, EDR and respiration. Different classifiers like k-NN, MLP, LDA, [SOFM](#sofm) (Self organizing feature map) and [QDC](#qda) (Quadratic discriminant classifier) were used out of which QDC performed the best. Apart from that, they also used [PCA](#pca) (principal component analysis) for dimensionality reduction. They found strong evidence that supported the use of features derived through non-linear methods. The data obtained through linear methods and reduced with PCA had low accuracies that ranged from 19% to 28% while features derived through non-linear methods showed an accuracy between 76% to 100% (2012). [Nogueira et al](#noguiera) applied a series of machine learning techniques to classify affective emotions hinged on the valence and arousal model and achieved high accuracies for all the classifiers. This could be owed to their data pre-processing methods. The features mainly consisted of HRV, skin conductance and EMG data and images were used as a mode of emotion elicitation. However, they noticed that mere images were not good enough to achieve this as the responses tend to be weak and the participants rate their emotions based in relative to the previous images instead of an unequivocal scale. The participants were asked to rate their emotions on a 10 point Likert scale and they were asked to listen to relaxing music and play a horror game before rating the images. The biosensor data was converted into polynomial values following which they used a correlation function instead of normalizing the data. They argued that normalizing the data results in an increase in noise due to the differences in individual experiences. The neural network consisted of ten neurons and achieved an accuracy of 97.5% which was the highest among all the other classifier. This was followed by SVM which used a linear kernel with a gamma values of 0.1 and 0.3, it wangled an accuracy of 95.65. They used a Laplacian error estimation function for decision trees to prevent overfitting which acquired an accuracy of 84% which was the lowest among all the four classifiers used. [Random forest](#randomforest) was supplemented with trees ranging between 50 to 5000 and was increased linearly, it attained an accuracy of 91.3% (2013).

# Proposed Solution:

As discussed in the [introduction](#_Introduction_and_Problem), the reliability of recruitment tests in determining the employability of a candidate is questionable. ([Rosse et al.](#rosse), 1998; [Miller and Barrett](#ref_4), 2008) The results of the personality tests could be easily corrupted either through response distortion or practice. We intend to study the viability of computational psychometrics at tackling the aforementioned problem. We will be monitoring biosensor data while our participants taking a test meant to measure the personality traits mentioned in the five factor model. This data will be classified with machine learning and finally we will be testing it against personality questionnaire to measure its reliability as an alternative. The five factor model has been accepted as a golden standard in personality research and we have seen some of the deeper implications of personality facets on an organization as a whole. ([Rothmann and Coetzer](#rothman), 2003) Personality has a considerable effect on an employee’s performance at workplace. ([Giberson et al.](#giberson), 2009; [Nadkarni and Herrmann](#nadkarni) , 2010; [Judge and Bono](#judge), 2000) CEO’s have a prominent influence on the performance of a company and its work culture. ([Costa and McCrae](#costa2008), 2008)We will be administering the NEO-PI-R personality inventory questionnaire on a computer. The NEO-PI-R consists of two forms i.e. S-form which is a self-report questionnaire and the R-form which is an observer report questionnaire. The participants will be asked to fill the S-Form. ([Watson and Clark](#watson), 1992; [Judge et al.](#judge_leadership) ,2002; [Panaccio et al](#panaccio).,2012) Affective emotions have a strong influence on personality traits. Positive valence has a strong influence on extraversion and conscientiousness and moderate effect on openness to experience and agreeableness while negative valence has a direct correlation with neuroticism. The participants will also be asked to pick a certain emotion in response to each question and rate it on a scale of five. This part would be based on the PANAS scale which is well accepted method for measuring affective states. These affective emotions will be classified as negative or positive and their ratings will be kept intact. This is to improve the accuracy of our classifier. This will be used as our class labels for machine learning later on. After reviewing pertinent literature, we can descry the efficiency of certain biosensors in obtaining information pertaining to emotional valence, stress and cognitive load to a reasonable but reliable extent. We intend to use HRV, EDR and skin temperature sensors to measure emotional arousal and valence. These sensors have an influence on autonomic nervous system which enable us to measure affective emotions. ([Appelhans and Luecken](#ref_8), 2006; [Hjortskov](#ref_9), 2004) A direct correlation has been established between emotion regulation and HRV through its influence on the parasympathetic nervous system. ([Thayer et al](#ref_10), 2009) HRV affects the vagus nerve which in turn affects the prefrontal cortex responsible for complex brain functions. ([Thayer et al](#ref_7), 2012) Apart from that, HRV also affects the amygdala which holds control over important operations like memory and emotion regulation. We can infer from all this data that HRV can be used as a biomarker for emotional wellbeing. It would also help us understand the changes in an individual's cognition in response to an event along with the ability to process a piece of data. ([Collet et al](#collet), 1997; [Villarejo, Zapirain, Zorrilla](#ref_13), 2012; [Kurniawan, Maslov, Pechenizkiy](#ref_14), 2013 ) Similarly experiments have been conducted to use EDR as a means to classify emotions and detect stress. The ability to measure changes in arousal has made it a relatively good at measuring and making a distinction between certain kind of emotions like anger and sadness. ([Shi et al](#ref_12), 2007) It can also be used to measure the changes in cognitive function.([Kaufman and Kagan](#ref_16), 1996) Variations in emotions also lead to changes in skin temperature. ([Lisetti and Nasoz](#ref_15),2004) It has been used to monitor changes in emotional states. ([Khan and Ingelby](#ref_17), 2009; [Puri et al](#ref_18), 2005; [Pavlidis, Levine, Baukol](#ref_19), 2000) Facial temperatures can help classify different emotional states with reliable accuracies. It can also detect acute stress and anxiety with a sustainable accuracy. ([Allison and Schnell](#ref_20), 2010) It can also be used to detect cognitive load in a fashion similar to the other sensors. We will be logging data from these three biosensors following which a threshold for the sensors will be set after the observing the data over a certain period of time. We will be using machine learning algorithms to classify the emotions to fit with the valence model. The biosensor data will act as the input variables while the PANAS ratings will act as the response variables for our training set. Different machine learning algorithms have been used previously to classify emotions using biosensor data with varying results out of which neural networks and algorithms based on neural networks, SVM an LDA have shown the most promise. ([Kim and Andre](#kimandre), 2008; [Jang et al](#jang), 2014) LDA generally showed the most consistency with high accuracies. However, this could be owed to the excess of features used in these studies. LDA works as a dimensionality reduction algorithm and we have a lack of features in our study which prevents us from using it. ([Yannakakis and Hallam](#yannakakis), 2007; [Lee et al](#lee_yoo), 2005; [Kim et al](#kimkim), 2004; [Nogueira et al](#noguiera), 2013; [Oliveira et al](#oliveira), 2011) SVM and algorithms based on neural networks have showed varying accuracies. They have performed better in comparison to other alternatives. However, the accuracies were still low on average. ([Lee et al](#lee_yoo), 2005; [Nogueira et al](#noguiera), 2013) Two studies manifested adequate results which leads us to believing that it is a matter of data pre-processing. We will be using a similar approach and convert the data to polynomial format and use a correlation function. ([Valenza et al](#valenza), 2012) Non-linear features performed very well in comparison to linear features. We will also consider this approach as an alternative. We will be using SVM due to its robustness and ability to customize kernels. Additionally, we would be using convolutional neural networks even though it’s primarily used for classifying images, it has shown promise at classifying biomedical data. We assume our accuracy would be much higher since we are only classifying it into two categories. We will be using the ratings we obtained for the classified and treating it as an additional response to the questionnaire by classifying it using with the same machine learning algorithms. The training set will consist of the actual answers to the NEO-PI-R personality inventory questionnaire. The sensor data will be sent directly to an IOT dashboard like ThingSpeak which comes with MATLAB analysis tools. The classified data will be compared and modelled against the original questionnaire.

# General Diagram of the proposed system

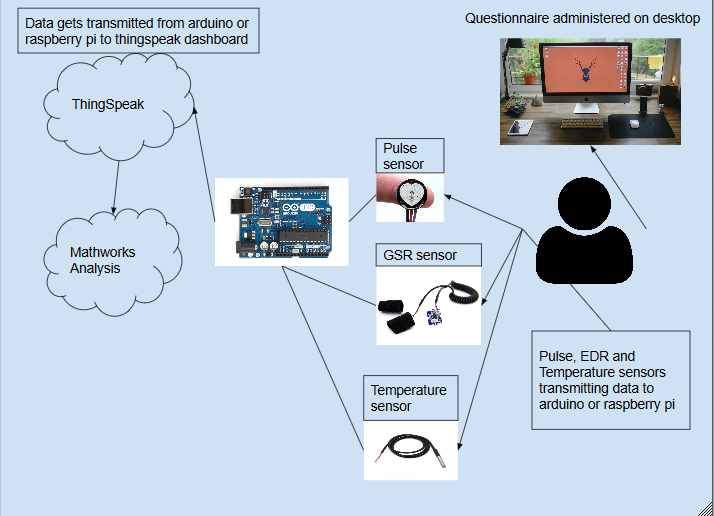


Figure Proposed System. In our system, the user will be hooked to Pulse, EDR(GSR) and temperature sensors while taking the NEO-PI-R and aptitutde tests. This data will be transferred to a dashobard like thingspeak where we may graph it and perform analysis.

# Methodology:

The research will proceed in four key phases.

### Phase 1: Development of the questionnaires

In this phase, we will review the relevant literature and find standardized questionnaires for cognitive and personality tests which are conducive to our experiment

### Phase 2: Setting up the IOT interface

We will begin by setting up the interface with MQTT protocol to transfer the data to an IOT dashboard like ThingSpeak by writing code for both ends. We will then develop a threshold mechanism for the sensors and divide it into five sections using the Likert scale.

### Phase 3: Preliminary test and analysis

In this phase, we will review the relevant literature and observe the data obtained from a test sample and check the reliability of our data and recalibrate system. The test sample will be classified with SVM and CNN to check its accuracy. A lower accuracy would lead to us considering other algorithms. This data will be used as a training set for our machine learning classifiers.

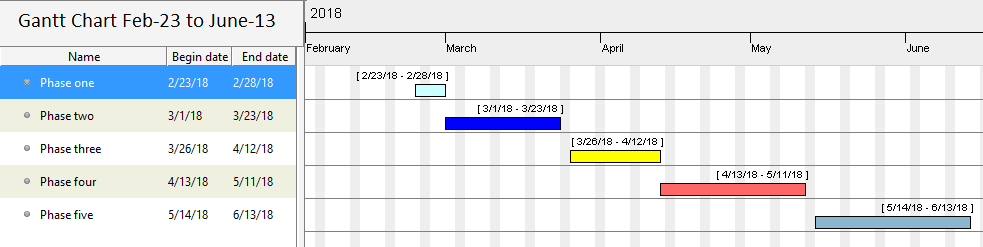
### Phase 4: Experiment

We will hold a mock interview process where the participants will take the psychometric tests while we monitor the biosensor data. The biosensor data will be classified into affective emotional rating following which it will be classified again into a Likert scale based on the personality questionnaire.

### Phase 5: Analysis

We will review pertinent literature and analyse our results.

# Gantt Chart:



# Methods of Data Collection and Analysis:

We will use quantitative research methods for our research. Our data will be collected through the personality and aptitude tests that will be administered on a computer. Additionally, the participants will be hooked with HRV, temperature and EDR sensors which will be transmitted to an IOT dashboard. The biosensor data will be divided into five categories similar to a Likert scale representing their emotional valence. This data will be classified with SVM and CNN and the biosensor data will be trained. The output variables will consist of ratings from the PANAS questionnaire which will be felt for each questionnaire. This data will be classified again but this time the output variables will consist of the actual output to the NEO-PI-R questionnaire. We will conduct a factor analysis by comparing the classified scores obtained through machine learning against the data obtained from the personality test. Additionally, we will calculate the [Cronbach’s alpha](#cronbach) to measure the internal consistency for the same. We may also conduct a multivariate analysis of variance (MANOVA) to compare the variance of the factors in a similar fashion.

# Technology:

We would require the following equipment to complete our research

* Arduino: 33-40$ / Raspberry Pi: 50-69$
* Pulse sensor: 25$
* DS18B20 temperature + breadboard + 4.7k ohms resistor: 19-20$
* Grove GSR sensor: 14$
* Jumper wires

# Ethical Issues:

This research will involve consenting adult human subjects. The participants will be required to take tests which will measure their aptitude and personality traits. We will also be collecting biological data like skin resistance, blood volume pulse and skin temperature from our participants by means of non-invasive biosensors which will be strapped to their non-dominant hand while they take the test. Subjects will be advised about the nature of the data being collected and their right to withdraw from the experiment should they feel the need to do so.

# Aims/Objectives of the project:

## 9.1 Purpose of the project:

We have pondered about the issues that come with the usage of recruitment tests that are so ubiquitous in nature. Our project tries to study the implications of having biosensors as an additional criterion for these tests. We believe that this might improve the reliability of such tests.

## Nature of the information sought:

We intend to do a quantitative analysis, so our data will be numerical. We would be using statistical and mathematical methods for our analysis.

## Research Questions:

* are affective emotions obtained through the means of biosensors enough to classify personality traits for the five-factor model?
* Can biosensor classification provide a similar or higher level of consistency than a traditional personality questionnaire?

# Outcomes/Outputs:

Improving the reliability of psychometric tests coupled with the usage of wireless biosensors could lead to organizations adopting it for their recruitment procedures. Finding a correlation between the personality traits and sundry of different intelligences with sensor data will have implications on cognitive sciences research.

# Limitations and Future work:

Portable pulse sensors are largely incapable of providing data with reliability on the same level as an ECG and respiration sensors require the user to breathe into a device which makes it uncomfortable resulting in data that maybe specious in nature. We believe that physical sensors are bound to reduce immersion and cause ambulatory difficulties, so we propose utilizing physical biosensors to find facial cues which could be then combined with Infrared thermography in the future for another study. Another alternative would involve applying the principles of ballistocardiography to obtain HRV. [Adib et al.](#adib) developed radio wave-based sensors capable of monitoring heart rate and respiration without any physical contact. Their sensors are capable of monitoring multiple people at the same time without being affected by movement or physical obstruction with relatively high accuracy and may act as a suitable alternative to physical sensors (2015). We also believe that our research will suffer from bias due to the study being cross-sectional. A largescale controlled study consisting of a diverse demographic can solve this problem. This research would act as a baseline for further research into this domain. Personality research classifying a wider array of affective emotions and measuring its influence on personality traits would improve the possibilities of such a system. Biological data combined with deep learning algorithms would bring out patterns and establish connections that were previously invisible.

# Glossary:

* abstract reasoning – *It is a measure of fluid intelligence, which is the ability to identify new patterns and apply them in a similar context.*

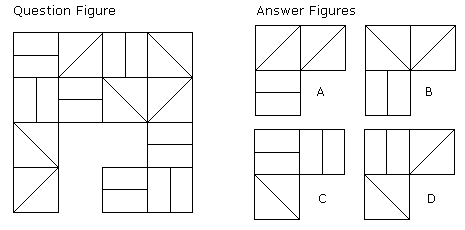


Figure . Abstract reasoning ([Psychometric Success](#ref_32), n.d). Solution to the problem is D as each row and column contains one of each type

* spatial reasoning –  *it is a measure of intelligence which is responsible for visualizing objects in three dimensions and deducing data with limited information.*

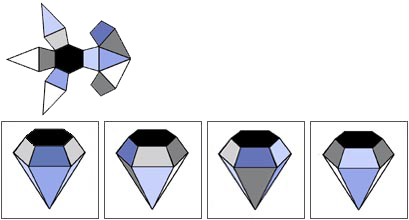


Figure . Spatial Reasoning([Fiboni](#ref_33),2017). Solution to this question is D.

* mechanical reasoning – *It is a measure of an individual’s ability to understand and applying mechanical principles.*

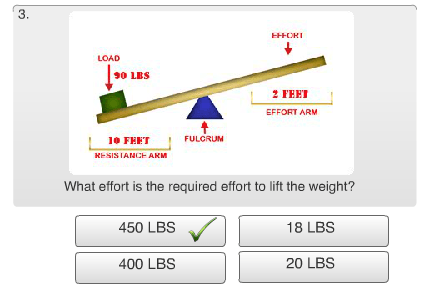


Figure . Mechanical Reasoning ([Steve S](#ref_34), 2013)

* cognition – *the process of knowledge acquisition performed by the brain. cognitive functions include memory, decision making, perception and reasoning.*
* autonomous nervous system – *A part of the peripheral nervous system that is responsible for the involuntary regulation of various bodily functions like digestion, pupil dilation, heartbeat, respiration, sneezing and so on.*
* sympathetic nervous system – *A part of the autonomic nervous system responsible for the fight or flight response. it increases blood flow, heart rate, blood pressure and respiration as a response to a stressful event. It is constantly active to maintain homeostatic functions like temperature regulation. Figure 5 describes the sympathetic nervous system.*
* parasympathetic nervous system *– The parasympathetic nervous system is complementary to the sympathetic nervous system and activates when the body is in a restful state. It is responsible for regulating certain digestive functions. Figure 5 describes the parasympathetic nervous system.*

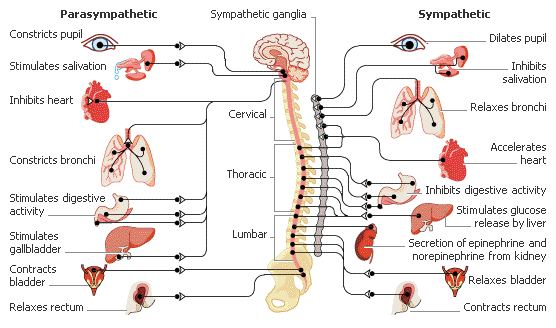


Figure . Sympathetic and Parasympathetic nervous system ([Medical Terms](#ref_35), 2014)

* amygdala *– It consists of two almond shaped set of neurons responsible for functions primarily responsible for emotional behaviour, decision making and generating fear in response to stimuli.*
* cardiovascular activity – *bodily functions Relating heart and blood vessels*
* vagus – *the vagus nerve is the longest cranial nerve and it is responsible for maintaining heart rate, respiration and digestion. it is also responsible for certain muscle movements.*
* cerebral – *pertaining to the brain.*
* situational awareness – *the ability of an individual to be aware of their surroundings, comprehend the events occurring within it and formulate an appropriate course of action.*
* working memory – *it is also referred to as short term memory as it is responsible for temporarily holding information.*
* machine learning – *A branch of computer science that deals with giving computers the ability to learn without being explicitly programmed.*
* electromyography – *A biomedical technology that is used to measure electrical activity generated by skeletal muscles.*
* human computer interface – *A field of computer science that observers the interaction between humans and computers.*
* Bayesian networks – *Bayesian networks are a form of probabilistic graphic model. It allows us to model the real world in the form of interconnected networks of probabilities. Bayesian networks cannot be cyclic which enables the algorithm to be updated faster. figure 6 represents a Bayesian network in which each circle represents a different world and the arrows indicate their relationship with each other. notice how the network is devoid of any cycles.*

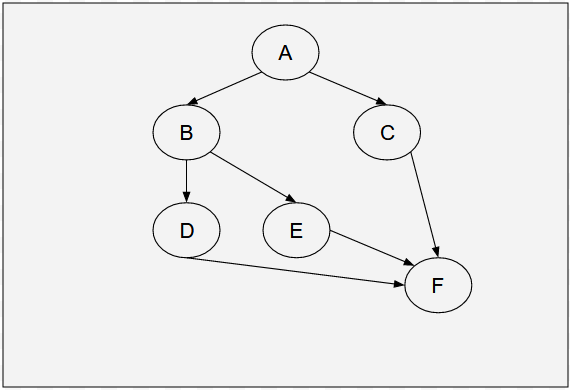


Figure . Bayesian networks

* decision trees – *decision trees is a form of tool that enables us to draw algorithms in the form of conditional loops. it starts off with a base nodes and progresses into multiple nodes based on the number of possible outcomes to a condition. figure 7 represents a simple decision tree for weather prediction. decision trees are also commonly used for machine learning, however, the conditions may be dependent on binary or decimal values of probabilities.*

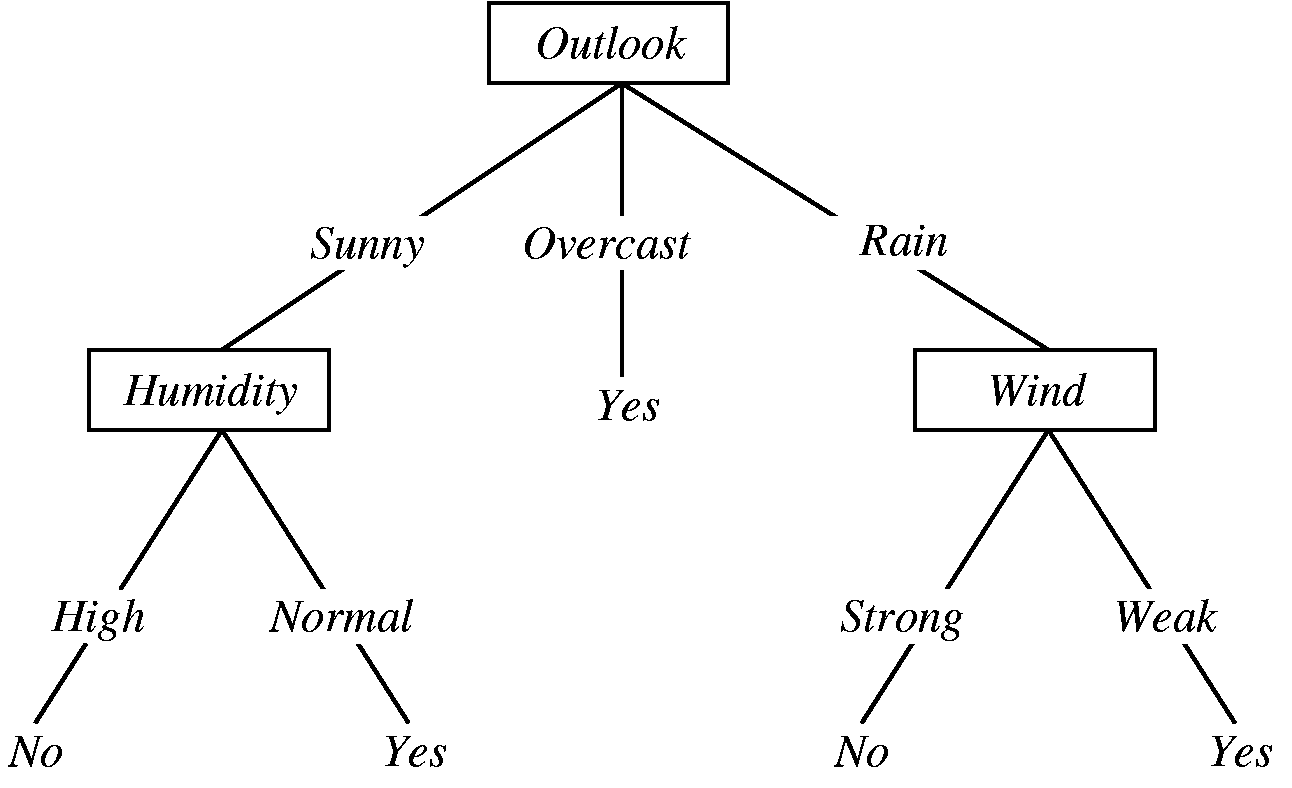


Figure . Decision tree ([Galea](#ref_36), 2016)

* sequential minimal optimization – smo is an optimization algorithm designed to solve the quadratic problem which appears with support vector machines. the quadratic problem is the problem of optimizing the quadratic function when several of the variables have linear constraints. this is done by minimizing or maximizing the quadratic function.
* stroop-word colour test *– it is a form of test designed to test the ability of an individual to inhibit the stroop phenomenon. this phenomenon is observed when there is an interference between two kinds of information. for example, if you try to say the colours in figure 8, you will find yourself reading the words instead.*

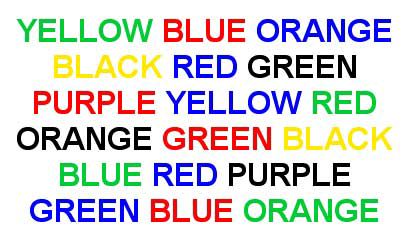


Figure . STROOP EFFECT ([Lorenzutti](#ref_37), 2012)

* Gaussian mixture model – *Gaussian mixture model is a form of unsupervised machine learning algorithm used for data clustering. it is a probabilistic function represented as a sum of multivariate Gaussian components with definite weight values. the weights are assigned based on its importance. this model is a super-position of normal distributions.*
* support vector machines *– support vector machines are supervised machine learning algorithms capable of solving classification and regression problems. consider a vector spaces with n dimensions representing consecutive features, svm will classify the data by finding the hyperplane (n-1 dimension) with the least bias between the features. in figure 9, the hyperplane passes right through the centre of the gap separating the trained features. this plane represents the largest distance between the features to prevent any bias.*

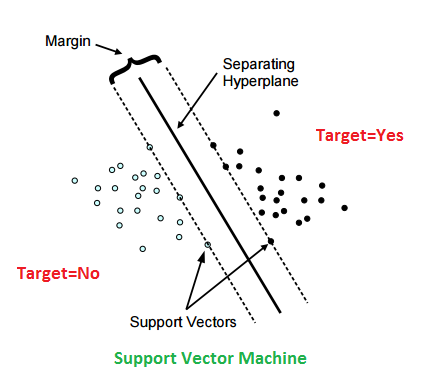


Figure . SVM ([DNI Institute](#ref_38), 2015)

* k-means clustering – *k-means clustering is a form of unsupervised machine learning algorithm used for data clustering. this algorithm works by defining k number of centroids and classifying the data into clusters by relating it with the nearest centroid. it is an iterative process and the centroids are initially defined randomly. the data is clustered in the consecutive iterations by associating it with the nearest mean. this process continues until a convergent point is achieved.*
* normalization *– it is the process of adjusting feature values to a common scale to prevent any probable disparity.*
* vasoconstriction – *vasoconstriction is the narrowing of blood vessels which results in an increase in blood pressure.*
* k-nearest neighbour algorithm – *k-NN is a classification algorithm used for both regression and classification problems. k-NN works by classifying averaging the values of its nearest neighbours.*
* discriminant function analysis – *it is a statistical method used to find the differences between continuous variables for multiple groups. it makes the assumption that the data is normally distributed.*
* marquardt’s backpropagation algorithm – *this is a machine learning algorithm developed to achieve 2nd order training speed without having to calculate scalar partial derivative functions which constitute a matrix known as the hessian matrix.*
* linear discriminant analysis – *it is a statistical and machine learning approach used to reduce dimensions in a feature vector by preserving the original differences between the classes. it assumes that the data is normally distributed and the labels are defined by the algorithm.*
* infrared thermography – *it is a technique used to obtain the thermal images of an object resulting from infrared radiation.*
* cognitive load – *it is a measure of cognitive effort which is observed when there is a strain on working memory.*
* artificial neural network – *it is a machine learning technique that is based on the neuronal network that compose an animal brain. neurons consist of axon and dendrites with terminals. signals are passed from dendrites to axon which in turn passes the signal to the dendrite of an another neuron. artificial neural networks work on the same principle, each network consists of an input layer defining the features and an output layer that defines the classification labels besides the many hidden layers which may or may not exist based on the complexity of the data. values passed onto the hidden layers result into the creation of new features. each layer will act as the activation of the consecutive layer.*
* steroid – *steroids are hormones responsible for regulating bodily functions like maintaining characteristics for each gender, signalling, regulating heartbeat and suppressing the immune system.*
* gluconeogenesis – *it is a series of chemical reactions that result in the formation of glucose from non-carbohydrate sources like proteins and amino acids.*
* interstitial fluid – *it is an extracellular fluid which can be found between tissue cells. close to 16% of the weight of human body is comprised of interstitial fluid. lymph, blood plasma and transcellular fluid (ocular, joint and cerebrospinal fluid) constitute to its components.*
* epilepsy *– a neurological disorder characterized by seizures that cause violent shaking with a surge of electrical activity within the brain.*
* adrenaline – *it is a hormone and a neurotransmitter produced by the adrenal gland in response to a fight or flight situation causing an increase in blood flow, blood sugar and heartbeat. it is induced by the adrenal medulla.*
* norepinephrine – *it is a hormone and a neurotransmitter produced by the adrenal gland and works in a similar fashion to adrenaline, however, it is generated throughout the day. it is induced by sympathetic nervous system.*
* brain computer interface – *brain computer interfaces (bci) which may also be referred to as brain machine interfaces or direct neural interface is a form of communication passage between a computing device and the brain established through the means of a collaboration of external hardware device and software extrapolating information from the brain.*
* short time Fourier transform – *it is an audio signal processing technique that maps the signal into a two dimensional functions consisting of time and frequency.*
* multilayer perceptron – *it is a form of artificial neural network that uses backpropagation (backward checking for error correction of the features in the network) and consists of one more layers situated between the input and output layers. the distinguishing feature of multilayer perceptron is its use of a non-linear activation function.*
* blood volume pulse – *it is an optical sensor used to measure the heart rate and heart rate variability.*
* virtual reality –  *a computer technology used to generate realistic virtual environments.*
* fuzzy systems – it is a control logic which acts on a true or false logic. instead of assuming the statements as absolute truths or falsehood, it attempts to break it into degrees of truth or falsehood.
* naive bayes – it is a form of classifier based on the Bayesian theorem which allows it to make assumptions based on probabilities. these probabilities are established after considering the previous results into consideration. it sees each feature as an independent entity.
* sofm – self organizing feature map is a form of unsupervised machine learning algorithm that is based on feed forward neural networks. it works as a dimensionality reduction algorithm and works well with a big feature set. it starts out with a big feature set and the learning rate reduces as it progresses to help the gradient descent converge.
* pca – principal component analysis is a dimensionality reduction algorithm. it works similar to linear discriminant analysis and transforms data through linear methods. however, it assigns the data to a different subset unlike lda which tries to determine the suitability of the data to a feature vector.
* qdc – quadratic discriminant classifier is a Bayesian classifier that uses a quadratic function. like most Bayesian classifiers, it treats each feature as an unrelated independent entity. it is considered to be an improvement over lda.
* random forests – it is an ensemble learning technique based on decision trees and tree bagging. it uses a random subset of values and builds multiple decision trees. random forests require large datasets to be efficient.
* cvi – *the competing values instrument is a scale used to measure the culture of an organization. it assumes that company culture is made of four values which are Clan, Adhocracy, Hierarchical and Market. it has the potential to unveil the intricacies of the management problem of an organization after measuring the existing culture and finding the values that may be desirable. a leadership model is then devised in the form of a quadrant with ratings to expose the areas that they need to focus more on.*
* panas – positive and negative affect schedule is a self-report questionnaire consisting binary ten item scales used to classify positive and negative emotionality.
* Cronbach’s alpha - *Cronbach’s alpha is used for measuring the internal consistency of a group of items. it does this by measuring how a set of items within a group are closely related.*
* sbs – sequential backward selector is a version of sequential feature selector used to classify data when there is a surplus of features. it is a dimensionality reduction algorithm and works by forming a subset of the best features that fit a criteria until it reduces the features to the desired amount.
* cnn – convolutional neural network is a form of deep neural network that is based on multilayer PERCEPTRON, but the input layer and hidden layer consists of convolutions which is the dot product of an n-dimensional matrix. the convolutions consist of filters that begin with basic patterns and take complex shapes as the network progresses.
* hlm – *hierarchical linear modelling is a statistical model based on the ordinary least square method (a statistical method used to obtain unknown values in a linear regression model) for hierarchical data. the data is generally in a nested format and it assumes that the data is normally distributed.*

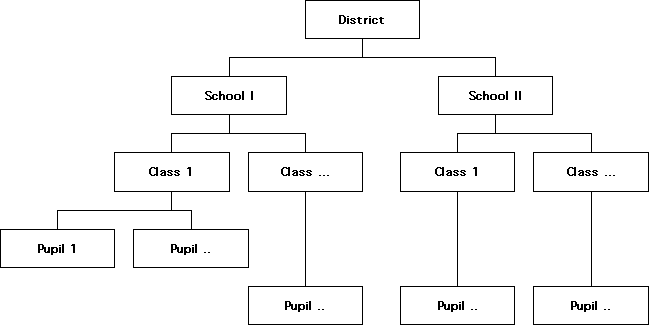


Figure HLM [(Pasdirtz](#HLM), 2011)

# References:

Hacker, C. (1997). The cost of poor hiring decisions ... and how to avoid them. HR Focus 74(10): 13-14.

Robert Half International. (2016, November 5). *The cost of a bad hire* [Press release]. Retrieved from <https://www.roberthalf.co.nz/press/cost-bad-hire>

Brady, J., McCormik, R., Crushell, B., Connor, C. O., Byrne, C., & Sheehan, C. (2017, November 12). What is Psychometric Assessment? Retrieved from http://aperturepartners.ie/index.php/2017/11/12/what-is-psychometric-assessment/

Donald, L. M. (2017, October 10). Axed Christchurch City Council staff could face controversial psychometric tests to reapply for jobs. *Stuff*. Retrieved from <https://www.stuff.co.nz/the-press/news/97726062/axed-christchurch-city-council-staff-could-face-controversial-psychometric-tests-to-reapply-for-jobs>

Miller, C. E., & Barrett, G. V. (2008). The Coachability and Fakability of Personality-Based Selection Tests Used for Police Selection. *Public Personnel Management,* *37*(3), 339-351. doi:10.1177/009102600803700306

Dakin, S., Nilakant, V., & Jensen, R. (1994). The Role of Personality Testing in Managerial Selection. *Journal of Managerial Psychology,* *9*(5), 3-11. doi:10.1108/02683949410066309

Cornforth, D. J., Tarvainen, M. P., & Jelinek, H. F. (2014). How to Calculate Renyi Entropy from Heart Rate Variability, and Why it Matters for Detecting Cardiac Autonomic Neuropathy. *Frontiers in Bioengineering and Biotechnology,* *2*. doi:10.3389/fbioe.2014.00034

Thayer, J. F., Åhs, F., Fredrikson, M., Sollers, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews,* *36*(2), 747-756. doi:10.1016/j.neubiorev.2011.11.009

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology,* *10*(3), 229-240. doi:10.1037/1089-2680.10.3.229

Hjortskov, N., Rissen, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Sogaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology,* *92*(1-2), 84-89. doi:10.1007/s00421-004-1055-z

Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart Rate Variability, Prefrontal Neural Function, and Cognitive Performance: The Neurovisceral Integration Perspective on Self-regulation, Adaptation, and Health. *Annals of Behavioral Medicine,* *37*(2), 141-153. doi:10.1007/s12160-009-9101-z

Iacono, W. G. (2001). Forensic “Lie Detection”. *Journal of Forensic Psychology Practice,* *1*(1), 75-86.

doi:10.1300/j158v01n01\_05

Shi, Y., Ruiz, N., Taib, R., Choi, E., & Chen, F. (2007). Galvanic skin response (GSR) as an index of cognitive load. *CHI 07 extended abstracts on Human factors in computing systems - CHI 07*. doi:10.1145/1240866.1241057

Villarejo, M. V., Zapirain, B. G., & Zorrilla, A. M. (2012). A Stress Sensor Based on Galvanic Skin Response (GSR) Controlled by ZigBee. *Sensors,* *12*(12), 6075-6101. doi:10.3390/s120506075

Kurniawan, H., Maslov, A. V., & Pechenizkiy, M. (2013). Stress detection from speech and Galvanic Skin Response signals. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*. doi:10.1109/cbms.2013.6627790

Collet, C., Vernet-Maury, E., Delhomme, G., & Dittmar, A. (1997). Autonomic nervous system response patterns specificity to basic emotions. *Journal of the Autonomic Nervous System,* *62*(1-2), 45-57. doi:10.1016/s0165-1838(96)00108-7

Lisetti, C. L., & Nasoz, F. (2004). Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *EURASIP Journal on Advances in Signal Processing,* *2004*(11), 1672-1687. doi:10.1155/s1110865704406192

Rimm-Kaufman, S. E., & Kagan, J. (1996). The psychological significance of changes in skin temperature. *Motivation and Emotion,* *20*(1), 63-78. doi:10.1007/bf02251007

Khan, M. M., Ward, R. D., & Ingleby, M. (2009). Classifying pretended and evoked facial expressions of positive and negative affective states using infrared measurement of skin temperature. *ACM Transactions on Applied Perception,* *6*(1), 1-22. doi:10.1145/1462055.1462061

Puri, C., Olson, L., Pavlidis, I., Levine, J., & Starren, J. (2005). StressCam. *CHI 05 extended abstracts on Human factors in computing systems - CHI 05*. doi:10.1145/1056808.1057007

Pavlidis, I., Levine, J., & Baukol, P. (2000). Thermal imaging for anxiety detection. *Proceedings IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (Cat. No.PR00640)*. doi:10.1109/cvbvs.2000.855255

Stemberger, J., Allison, R. S., & Schnell, T. (2010). Thermal Imaging as a Way to Classify Cognitive Workload. *2010 Canadian Conference on Computer and Robot Vision*. doi:10.1109/crv.2010.37

Venugopal, M., Arya, S. K., Chornokur, G., & Bhansali, S. (2011). A realtime and continuous assessment of cortisol in ISF using electrochemical impedance spectroscopy. *Sensors and Actuators A: Physical,* *172*(1), 154-160. doi:10.1016/j.sna.2011.04.028

Lumicao, M. N., Levendowski, D. J., Berka, C., Yau, A., Davis, G., Zivkovic, V. T., . . . Craven, P. L. (june 2007). Eeg correlates of task engagement and mental workload in vigilance learning and memory tasks. *Aviation Space and Environmental Medicine,* *78*(B231), 44th ser., 187-196.

Lin, Y., Wang, C., Jung, T., Wu, T., Jeng, S., Duann, J., & Chen, J. (2010). EEG-Based Emotion Recognition in Music Listening. *IEEE Transactions on Biomedical Engineering,* *57*(7), 1798-1806. doi:10.1109/tbme.2010.2048568

Hou, X., Liu, Y., Sourina, O., & Mueller-Wittig, W. (2015). CogniMeter: EEG-based Emotion, Mental Workload and Stress Visual Monitoring. *2015 International Conference on Cyberworlds (CW)*. doi:10.1109/cw.2015.58

Cipresso, P. (2015). Modeling behavior dynamics using computational psychometrics within virtual worlds. *Frontiers in Psychology,* *6*. doi:10.3389/fpsyg.2015.01725

Villani, D., Cipresso, P., Repetto, C., Bosone, L., Balgera, A., Mauri, M., . . . Riva, G. (2015). Computational Psychometrics in Communication and Implications in Decision Making. *Computational and Mathematical Methods in Medicine,* *2015*, 1-10. doi:10.1155/2015/985032

Segkouli, S., Paliokas, I., Tzovaras, D., Tsakiris, T., Tsolaki, M., & Karagiannidis, C. (2015). Novel Virtual User Models of Mild Cognitive Impairment for Simulating Dementia. *Computational and Mathematical Methods in Medicine,* *2015*, 1-15. doi:10.1155/2015/358638

Davier, A. A. (2017). Computational Psychometrics in Support of Collaborative Educational Assessments. *Journal of Educational Measurement,* *54*(1), 3-11. doi:10.1111/jedm.12129

Costa, P. T., & Mccrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology,* *54*(5), 853-863. doi:10.1037/0022-3514.54.5.853

Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology,* *41*(1), 417-440. doi:10.1146/annurev.ps.41.020190.002221

Mccrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology,* *52*(6), 1258-1265. doi:10.1037/0022-3514.52.6.1258

Kausch, O., Rugle, L., & Rowland, D. Y. (2006). Lifetime Histories of Trauma among Pathological Gamblers. *American Journal on Addictions,* *15*(1), 35-43. doi:10.1080/10550490500419045

Smits, S. J., Mclean, E. R., & Tanner, J. R. (1993). Managing High-Achieving Information Systems Professionals. *Journal of Management Information Systems,* *9*(4), 103-120. doi:10.1080/07421222.1993.11517981

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences,* *29*(6), 1057-1068. doi:10.1016/s0191-8869(99)00253-6

Graziano, W. G., Jensen-Campbell, L. A., & Hair, E. C. (1996). Perceiving interpersonal conflict and reacting to it: The case for agreeableness. *Journal of Personality and Social Psychology,* *70*(4), 820-835. doi:10.1037/0022-3514.70.4.820

Rothmann, S., & Coetzer, E. P. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology,* *29*(1). doi:10.4102/sajip.v29i1.88

Giberson, T. R., Resick, C. J., Dickson, M. W., Mitchelson, J. K., Randall, K. R., & Clark, M. A. (2009). Leadership and Organizational Culture: Linking CEO Characteristics to Cultural Values. *Journal of Business and Psychology,* *24*(2), 123-137. doi:10.1007/s10869-009-9109-1

Cheung, K., & Keeves, J. (1990). Hierarchical linear modelling. *International Journal of Educational Research,* *14*(3), 289-297. doi:10.1016/0883-0355(90)90039-b

Nadkarni, S., & Herrmann, P. (2010). CEO Personality, Strategic Flexibility, and Firm Performance: The Case of the Indian Business Process Outsourcing Industry. *Academy of Management Journal,* *53*(5), 1050-1073. doi:10.5465/amj.2010.54533196

Booth, T., Murray, A. L., Overduin, M., Matthews, M., & Furnham, A. (2015). Distinguishing CEOs from Top Level Management: A Profile Analysis of Individual Differences, Career Paths and Demographics. *Journal of Business and Psychology,* *31*(2), 205-216. doi:10.1007/s10869-015-9416-7

Judge, T. A., & Bono, J. E. (2000). Five-factor model of personality and transformational leadership. *Journal of Applied Psychology,* *85*(5), 751-765. doi:10.1037//0021-9010.85.5.751

Palaiou, K., & Furnham, A. (2014). Are bosses unique? Personality facet differences between CEOs and staff in five work sectors. *Consulting Psychology Journal: Practice and Research,* *66*(3), 173-196. doi:10.1037/cpb0000010

Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology,* *83*(4), 634-644. doi:10.1037//0021-9010.83.4.634

Watson, D., & Clark, L. A. (1992). On Traits and Temperament: General and Specific Factors of Emotional Experience and Their Relation to the Five-Factor Model. *Journal of Personality,* *60*(2), 441-476. doi:10.1111/j.1467-6494.1992.tb00980.x

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology,* *87*(4), 765-780. doi:10.1037//0021-9010.87.4.765

Panaccio, A., & Vandenberghe, C. (2012). Five-factor model of personality and organizational commitment: The mediating role of positive and negative affective states. *Journal of Vocational Behavior,* *80*(3), 647-658. doi:10.1016/j.jvb.2012.03.002

Mandryk, R. L., & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies,* *65*(4), 329-347. doi:10.1016/j.ijhcs.2006.11.011

Yannakakis, G. N., & Hallam, J. (n.d.). Entertainment Modeling in Physical Play Through Physiology Beyond Heart-Rate. *Affective Computing and Intelligent Interaction Lecture Notes in Computer Science,* 254-265. doi:10.1007/978-3-540-74889-2\_23

Chanel, G., Rebetez, C., Bétrancourt, M., & Pun, T. (2008). Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era - MindTrek 08*. doi:10.1145/1457199.1457203

Holehouse, A. (2012, January 9). Introduction, Regression Analysis, Gradient Descent [Regression]. Retrieved from http://www.holehouse.org/mlclass/01\_02\_Introduction\_regression\_analysis\_and\_gr\_files/Image.png

Lee, C., Yoo, S., Park, Y., Kim, N., Jeong, K., & Lee, B. (2005). Using Neural Network to Recognize Human Emotions from Heart Rate Variability and Skin Resistance. *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. doi:10.1109/iembs.2005.1615734

Kim, J., & Andre, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* *30*(12), 2067-2083. doi:10.1109/tpami.2008.26

Costa, P. T., & Mccrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology,* *54*(5), 853-863. doi:10.1037/0022-3514.54.5.853

Ferdinando, H., Ye, L., Seppanen, T., & Alasaarela, E. (2014). Motion Recognition by Heart Rate Variability . *Australian Journal of Basic and Applied Sciences ,* *ISSN*(1991-8178), 50-55. Retrieved from https://www.researchgate.net/profile/Hany\_Ferdinando/publication/267212566\_Emotion\_Recognition\_by\_Heart\_Rate\_Variability/links/57dfa8f308aeea19593b693a/Emotion-Recognition-by-Heart-Rate-Variability.pdf?origin=publication\_detail.

Kim, K. H., Bang, S. W., & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical & Biological Engineering & Computing,* *42*(3), 419-427. doi:10.1007/bf02344719

Oliveira, E., Benovoy, M., Ribeiro, N., & Chambel, T. (2011). Towards Emotional Interaction: Using Movies to Automatically Learn Users’ Emotional States. *Human-Computer Interaction – INTERACT 2011 Lecture Notes in Computer Science,* 152-161. doi:10.1007/978-3-642-23774-4\_15

Martinez, H. P., Bengio, Y., & Yannakakis, G. N. (2013). Learning deep physiological models of affect. *IEEE Computational Intelligence Magazine,* *8*(2), 20-33. doi:10.1109/mci.2013.2247823

Jang, E., Park, B., Kim, S., Chung, M., Park, M., & Sohn, J. (2014). Emotion classification based on bio-signals emotion recognition using machine learning algorithms. *2014 International Conference on Information Science, Electronics and Electrical Engineering*. doi:10.1109/infoseee.2014.6946144

Valenza, G., Lanata, A., & Scilingo, E. P. (2012). The Role of Nonlinear Dynamics in Affective Valence and Arousal Recognition. *IEEE Transactions on Affective Computing,* *3*(2), 237-249. doi:10.1109/t-affc.2011.30

Nogueira, P. A., Rodrigues, R., Oliveira, E., & Nacke, L. E. (2013). A Hybrid Approach at Emotional State Detection: Merging Theoretical Models of Emotion with Data-Driven Statistical Classifiers. *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. doi:10.1109/wi-iat.2013.117

Costa, P. T., & Mccrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). *The SAGE Handbook of Personality Theory and Assessment: Volume 2 — Personality Measurement and Testing,* 179-198. doi:10.4135/9781849200479.n9

[Abstract Reasoning]. (n.d.). Retrieved from http://www.psychometric-success.com/images/AA0602.gif

Fiboni, V. O. (2017). [Spatial Reasoning]. Retrieved from https://www.fibonicci.com/images/spatial/test-spatial-hard-10.png

S, S. (2013). [Mechanical Reasoning]. Retrieved from https://mechanicalaptitudetest.org/wp-content/uploads/2013/01/R1Q3.png

[Sympathetic and Parasympathetic nervous system]. (2014). Retrieved from http://medicalterms.info/img/uploads/anatomy/autonomic-nervous-system.gif

Galea, M. (2016, August 20). [Decision Tree]. Retrieved from <http://cloudmark.github.io/images/kotlin/ID3.png>

Lorenzutti, C. (2012, March 12). [STROOP effect]. Retrieved from https://lh6.googleusercontent.com/-m1Hhxr3WwkE/T19TGgTg\_bI/AAAAAAAAEnY/c0xeh6R-jO0/w506-h750/fig\_20\_stroop\_test.jpg

DNI institute. (2015, September 13). [Support Vector Machines]. Retrieved from http://dni-institute.in/blogs/wp-content/uploads/2015/09/SVM-Planes.png

Pasdirtz, G. W. (2011, December 16). [Hierarchical linear modelling]. Retrieved from <http://1.bp.blogspot.com/-XDU3NyXhSGw/TuupxhDZSPI/AAAAAAAABF8/FBZr5ENiDi8/s1600/Hierarchical%2BSchool%2BSystem.GIF>

Adib, F., Mao, H., Kabelac, Z., Katabi, D., & Miller, R. C. (2015). Smart Homes that Monitor Breathing and Heart Rate. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI 15*. doi:10.1145/2702123.2702200