myCAST: A Personalized Stroke Identification and Prevention System

Swagat Bhattacharyya* and Michael Yan*

*Senior, Morgantown High School, Morgantown, WV 26501

***Abstract*—Strokes are a leading cause of death throughout the world, yet little has been done to assess stroke susceptibility or provide retainable education on stroke prevention. Limited health care coverage in many areas coupled with the presumed unlikeliness of stroke occurrences may cause many to disregard symptoms, delaying treatment and thereby reducing survival rates. In fact, victims hospitalized within three hours of their first symptoms display drastically increased survival rates. Hence, the concept of a mobile app for stroke self-examination and education was explored. Self-examination was performed using a three-stage procedure. The first stage of self-examination was a questionnaire that tested the user's sensations of moderate stroke indicators. Detection of slurred speech, a prominent stroke sign, followed. Users were prompted to repeat a predetermined phrase, which was converted to text using Google's Speech-to-Text API. This text was compared to the original prompt using a string distance metric. The final stage was a face paresis detector developed in MATLAB®. The detector captured a smiling self-portrait of the user. After identification, the user's mouth profile was categorized into an asymmetry criterion which could be compared to archived data to yield another stroke identification metric. Finally, the proposed mobile app compiled the various identification metrics and alerted the user to seek medical assistance accordingly. An innovative emergency contact button sent the user's GPS location and medical history to emergency services, fostering expedient communication and treatment. The mobile app was validated on the authors (inputting American Stroke Association identified stroke symptoms).**

Keywords*—*Early Stroke Identification, Stroke Risk Assessment, Stroke Education, Facial Analysis, Speech Slur Analysis, Telemedicine

I. Introduction

A. Background

Strokes were the second leading cause of death worldwide in 2015, leading to over 6.24 million fatalities [1]. However, little has been done to assess stroke susceptibility in potential victims. In addition, public awareness on the subject matter is lacking, although education of those at risk provides a simple way to prevent strokes and accelerate emergency response. In fact, a 2005 survey found that over 62% of American adults could not correctly identify the major stroke symptoms [2]. Stroke's sparse nature renders it infeasible for potential victims to get frequent physician checkups — especially in regions with geographic isolation or in cases of limited health care coverage. Geographic isolation also slows emergency response — shortening the time interval for treatment. Concern about the ever-increasing medical fees (even in the absence of stroke), limited health-care coverage, and the presumed unlikeliness of a stroke cause many at risk to disregard symptoms — delaying treatment and reducing survival rates. In contrast, if early-onset stroke signs are detected and victims rushed to hospitals within three hours, survival rates would dramatically increase [2].

There are two major stroke categories. Ischemic strokes occur the most frequently, comprising 87% of all stroke cases, and result from blood vessel obstructions [3]. Some ischemic strokes can be predicted in advance from the occurrence of Transient Ischemic Attacks (TIAs), which are caused by temporary blood flow blockages. TIAs have symptoms similar to that of ischemic strokes but not the same long-lasting effects on the brain [4]. TIAs can potentially predict the onset of a future stroke, and their identification could help accelerate stroke treatment. Alteplase IV r-tPA, the "gold standard" treatment for ischemic strokes, is only effective in an acute three-hour window after stroke onset [5], elucidating the need for fast identification and treatment.

The other major stroke type, hemorrhagic, occurs in only 15% of all stroke cases but has greater fatality rate. Hemorrhagic strokes are caused by leakage from a weakened blood vessel or the burst of a brain aneurysm [6]. Treatment for hemorrhagic strokes typically involves prompt surgery [5]. However, in certain instances stroke victims do not experience symptoms. Such a stroke is dubbed a "silent stroke" and may lead to long term damage primarily due to victims not receiving treatment [3]. The proposed mobile app was targeted toward the detection of ischemic strokes, hemorrhagic strokes, and TIAs. Determination of long-term trends could aid in the detection of "silent" strokes.

B. Objectives and Novelty

The prevalence of mobile devices globally provides a promising solution for early detection of stroke. Furthermore, mobile devices have little operational cost and offer the user great convenience. Advancements in mobile sensing and processing made it possible to leverage these advantages to develop Cerebrovascular Accident Self-Test (myCAST), the mobile app proposed in this paper. myCAST was designed to be a general-purpose tool and was hence built as a three-layer system.

The first layer provided education for stroke prevention. The stroke prevention provision was intended to construct an educational environment in order to teach users about stroke symptoms and prevention. The proposed system had the potential of enhancing information retention.

The second layer performed an automated self-examination based on user reported symptoms and analysis of exhibited symptoms. A cascade of three mini-tests was utilized within this self-examination. The first mini-test was a questionnaire that inquired of the user's perception of moderate stroke indicators. Detection of slurred speech, a prominent stroke sign [7], followed. The final mini-test was a facial paresis detection scheme. The results of the aforementioned mini-tests were compiled into a weighted composite score. myCAST alerted the user to seek medical assistance based on the score magnitude. The rigor of the final two mini-tests give myCAST the potential to achieve higher accuracies and differentiate from the applications currently on the market. Additionally, the testing frequency and the final score are user-specific, adding an immense level of reliability.

The purpose of the final layer was to expedite emergency communication. Key to this was an intelligent, emergency-contact button that sent the user's GPS location and medical history to emergency workers if stroke symptoms were strongly exhibited. myCAST achieved this through available mobile technology, such as Google Text-to-Speech API, and through SMS to the Federal Communication Commission's (FCC's) growing "Text-to-911" program. In this way, myCAST could act as an expedient link between users, dispatchers, and emergency medical technicians (EMTs), significantly reducing communication overhead.

C. Overview of Paper

This paper is structured as follows. Section II discusses work in related areas and contrasts their benefits and drawbacks. Section III presents the inner workings of the proposed mobile application in depth. The performance evaluation of the application and possibilities for future work are discussed in Section IV. Section V provides a summary of the paper.

II. Related Work

A. Previous Mobile Applications

Historically, the scope of medical self-diagnosis mobile apps has been limited. The vast majority of mobile apps pertaining to medical topics are informational pamphlets, usually offering vague, impersonal, risk factors and symptoms. Those that do attempt self-examinations are usually detached from physician support and analyze only rudimentary symptoms, resulting in low confidence intervals for stroke detection and risk assessment. Stroke Riskometer [8] is a $3.00 app on the Google Play Store that attempts a new approach to self-risk assessment. The app consists of a survey, and offers a broad "5-year stroke risk prediction" metric based on survey answers and age/sex indicators. In general, however, the software currently available for stroke self-examination is overgeneralized, for-profit, and imprecise.

B. Immersive Learning

The recognition of a given scenario in daily life actually refers to a person's recollection ability of their recognition memory. Recognition memory is a category of declarative memory and is defined as the ability to recognized previously experienced (or simulated) events [9]. Since most stroke victims do not recognize their symptoms as those of stroke (typically due to stroke's apparent "unlikeliness" and overgeneralizable symptoms) an educational environment that provides visuo-symbolic simulations of major symptoms would aid users to encode these experiences into recognition memory. In the long run, this would be effective in helping familiarity recognition and the recall of state-dependent learning/memory, thereby increasing the rate at which first-time victims recognize previously unexperienced stroke symptoms. In order to achieve maximum memory retention, interactive, visually appealing [10], and representationally accurate environments should be provided.

C. Slurred Speech

One of the most prominent indicators of stroke is the presence of slurred or garbled speech [11]. In order to differentiate this form of speech from non-impeded speech, numerous algorithms analyzing raw audio feed have been researched and developed, all with limited success in a mobile environment. These algorithms also use computationally intensive transformations, such as the Fourier transform, to dissect speech input and compare it to databases of normal words. Due to their computational rigor, integration of these algorithms into mobile platforms is often infeasible for routine operation.

D. Face Asymmetry Detection

Facial asymmetry is a natural consequence of facial paresis — a major indicator of stroke onset. However, a gradual onset or general lack of self-awareness may render the victim unaware of such a circumstance. Historically, facial recognition software on the mobile platform has been extensively used; however, most of this use was to merely identify the presence of faces in the camera feed by utilizing cascade object detection frameworks [12]. Traditionally used image processing techniques, such as the Principal Component Analysis [13], require intense mathematical computation, rendering them infeasible on the mobile platform.

III. Methods and Algorithms

A. Education

The education tab is the second provision of the mobile app and showcases common stroke symptoms in an easily accessible format. The major stroke symptoms as popularized by the American Stroke Association's "FAST" acronym [3] are featured in this tab. Other major symptoms are listed alongside relevant symbols. Inclusion of symbols and acronyms are speculated to improve associative memory in users and hence accelerate information retrieval. A screen capture of myCAST's home screen and education tab are displayed in Figure 1 and Figure 2, respectively.

B. Questionnaire

In order to determine the user's perception of moderate stroke indicators [7] that are physically difficult to measure, a short questionnaire is presented to the user at the start of the self-examination. This questionnaire is short and easily readable in order to be able to be completable by the potentially disoriented or impaired user. The following questions are asked in the questionnaire (the process used to determine which questions to ask is described in more detail in Section III. E):

1. Have you recently experienced blurred or tunnel vision?
2. Have you recently felt any weakness on one side of your body or face?
3. Have you recently felt SUDDEN confusion, dizziness, or nausea?

Since the symptoms described have little probability of occurring on a regular basis there is little need for archiving questionnaire responses. Answers are simply stored in a temporary array for future reference. Figure 3 exhibits a screen capture of the questionnaire.

C. Speech Slur Detection

Slurred speech is a key stroke indicator. In order to detect this symptom, which may appear promptly after stroke onset [7], myCAST prompts users to repeat a predetermined phrase. The resultant microphone feed is converted to text via Google's Speech-to-Text API. This text is then compared to the text in the prompt using the Levenshtein string distance metric [14]. The Levenshtein Distance between two strings is given by:

$$lev_{a,b}(i,j) = \begin{cases} 0, & i = j = 0 \\ i, & j = 0 \wedge i > 0 \\ j, & i = 0 \wedge j > 0 \\ L, & otherwise \end{cases} \tag{1}$$

$$where \quad L = min \begin{cases} lev_{a,b}(i-1,j) + 1 \\ lev_{a,b}(i,j-1) + 1 \\ lev_{a,b}(i-1,j-1) + \varphi_{i,j} \end{cases} \tag{2}$$

$$and \quad \varphi_{i,j} = \begin{cases} 1, & a_i \neq b_j \\ 0, & a_i = b_j \end{cases} \tag{3}$$

In Equation (3), a and b denote the two strings to be compared. Equation (1) yields an integer which can either be appended to a database (as is typical during training) or be compared to training values using a modified form of Pearson's χ^2 test for independence [15] given by:

$$J \frac{1}{n} \sum_{i=1}^{n} \left(\frac{(T_i - C)^2}{T_i}\right) \tag{4}$$

In Equation (4), n, T, and C denote the total number of training samples, the Levenshtein Distances of the training samples, and the current value of the Levenshtein Distance. The modified χ^2 test allows the algorithm to differentiate between slurred speech and innate speech impediment, resulting in a more accurate determination of slurred speech given an appropriately sized database of user samples. The modified χ^2 statistics and raw Levenshtein Distances obtained from this mini-test are stored in an internal SQLite database [16] for future retrieval and use in comparison tests. A screen capture of the speech slur mini-test is presented in Figure 4.

D. Facial Paresis Detection

Paresis of facial muscles is another key indicator of strokes [7]. Facial paresis is often characterized by an asymmetric smile. Hence, in order to detect facial paresis, the asymmetry in the user's smile is assessed by testing the fit of a quadratic model. First, myCAST prompts the user to capture a smiling self-portrait. Localized image processing follows. This processing was intended to be computationally minimalistic due to resource constraints on mobile devices. After capture, the user's smiling self-portrait is converted into grayscale. The contrast of this grayscale image is adjusted by 1% to mitigate the effects of low lighting, which increases the failure rate of the Haar-like feature detector [17] in the facial recognition stage. Though this step increases the effect of glare, it is anticipated that the application will be used in an indoor environment, where excessive lighting is typically not an issue. Facial recognition is performed next; upright faces are located in the grayscale image using the Viola-Jones algorithm [17].

The Viola-Jones algorithm has nearly perfect detection accuracy but frequently false triggers. In cases of multiple detections, the most prominent detection (defined as the face whose bounding box occupies the most area) is chosen for further analysis (i.e. the image is cropped at the largest bounding box's location). After that, the Viola-Jones algorithm is used to detect the user's mouth in the previously found facial region. The more constrained search region aids in determination of the true mouth location. From stock images, it is observed that the true mouth region typically satisfies the following two properties due to human construction:

1. Had an upper bound lower than 70% of the total height of the facial frame.
2. Was at least 15% of the facial frame in width.

Nevertheless, even after application of the aforementioned selection criteria, multiple mouth bounding boxes remained in the stock photos. The likeliness that a detection is actually a false trigger decreases with increasing size of the Haar-like features; thus, it can be inferred that the bounding box with the largest area (that also complies with the aforementioned criteria) is most likely to dictate the true mouth region of the user. After the algorithm selects an appropriate bounding box, it crops the facial frame to encompass only the location specified by the mouth bounding box.

Bounding box adjustment was considered in order to better capture the user's mouth profile. This method sought to reduce the influence of extraneous factors such as facial hair or an off-center bounding box on a facial asymmetry metric. Adjustment was accomplished using the MATLAB®

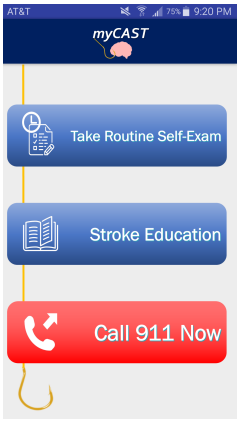

Fig. 1: Home screen

Fig. 2: Education tab

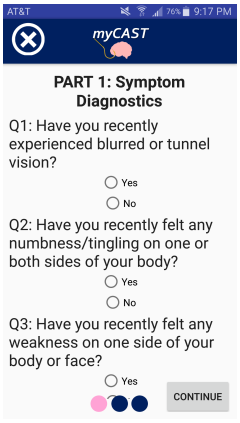

Fig. 3: Questionnaire

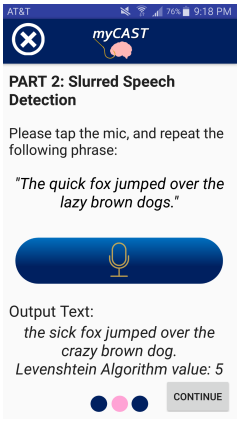

Fig. 4: Speech slur test

"fmincon" optimizer [18]. The optimizer solved a constrained nonlinear programming problem using a sequential quadratic programming (SQP) method [19]. During each iteration, the optimizer executed a function that first extracted and binarized the mouth profile at the region specified by the given bounding box parameters using an adaptive thresholding scheme [20] and then used k-means clustering [21] to categorize the pixels in the binarized outline into two clusters: the dominant (presumably the one containing the true mouth profile) cluster and a cluster that contained the outliers. The optimization problem solved was given by:

$$\arg\min_{U} -\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{5}$$

$$s.t.\ LB \leq U \leq UB \tag{6}$$

In Equation (5), x_2, x_1, y_2, and y_1 denote the Cartesian coordinates of the centroids of the two clusters, respectively. In Equation (6), LB, UB, and U denote the lower bound of the bounding box parameters, the upper bound of the bounding box parameters, and the best bounding box parameters, respectively. Although bounding box adjustment was often observed to yield a more robust fit of the mouth, it is counterproductive; the algorithm performs a considerable amount of computation, impacting algorithm speed significantly, while giving only a marginally better end result. Hence, the NP-hard bounding box correction is not used in the process of mouth profile extraction. Instead, a more robust asymmetry metric is pursued.

The next layer of processing is standardization, which is performed to account for varying camera-mouth distances and different user profiles. Scaling is based on the radii and centroid locations of regions present in the image. First, contiguous regions with activated pixels are found in the binarized image using the MATLAB® "regionprops" function [22]. This function computes the axes lengths corresponding to each region by constructing an ellipse with the same "moment of inertia" [23]. Axes lengths are averaged and halved to approximate the region radii. The largest radius, which theoretically corresponds to the region encased within the confines of the user's lip outline is used to scale the image such that the region occupies a 100-pixel radius (deemed a good scaling factor by visual observation. Bicubic interpolation is used to fill missing pixels after resizing [24]. The centroid location corresponding to the aforementioned region in the resized mouth profile is obtained, and the image is cropped to occupy 125 pixels in height (also found via visual observation) centered around it. This is done in order to remove extraneous points, which could otherwise bias the asymmetry criterion (based off the goodness of fit of a quadratic model).

Mouth outline tilt is then corrected and the profile is iteratively cleaned in order to further remove bias from the asymmetry criterion. First, the median location of the activated pixels in each column of the mouth profile and their corresponding weightages are determined to obtain numeric data. Not all columns have activated pixels; only columns with activated pixels were used henceforth for computational ease. The column number corresponding to each median value is also stored. A weight vector for the median location vector is created to record variances and is given by:

$$\varphi_i = \begin{cases} \frac{1}{\sigma_i^2}, & \sigma_i \neq 0 \\ 1, & \sigma_i = 0 \end{cases} \tag{7}$$

In Equation (7), σ denotes the standard deviation of the row numbers of the activated pixels in column i of the mouth profile. Then, to reduce noise in the previously found median location data, these data are filtered using a moving average filter. The Z-domain transfer function for a moving average filter [25] is given by:

$$\frac{Y(z)}{M(z)} = \frac{1}{m} \sum_{i=0}^{m-1} z^{-i} \tag{8}$$

In Equation (8), m and $M(z)$ denote the window size of the moving average filter and the discrete-time Laplace transform

of the median location data. A window size of ten pixels seemed to work well for the algorithm presented. The first m values are undefined and thus removed from the filtered median data and the weight vector. The remaining weights are compiled into a diagonal weight matrix given by:

$$\mu = diag(\varphi_1, ..., \varphi_n) \tag{9}$$

In Equation (9), n denotes the size of the weight vector found in Equation (7). The weight matrix is utilized to fit a first order polynomial model to the median data using the method of weighted linear least-squares. The linear model can be represented in the matrix form:

$$y = \Psi H;\ \Psi = \begin{bmatrix} X_1 & 1 \\ \vdots & \vdots \\ X_n & 1 \end{bmatrix} \tag{10}$$

In Equation (10), X and H denote the column number vector and the linear parameter vector, respectively. The weighted linear least-squares estimate [26] for parameter vector $\hat{H}$ is given by:

$$\hat{H} = (\Psi^T \mu \Psi)^{-1} \Psi^T \mu Y \tag{11}$$

In Equation (11), Y denotes the aforementioned median location vector. The value of $\hat{H}_1$, the slope of the linear model, is used to correct the tilt of the mouth profile. The original tilt of the mouth profile was approximated by:

$$\theta \approx \tan^{-1} \hat{H}_1 \tag{12}$$

Hence, the original coordinates of the median locations are represented as a rotation of a tilt-corrected set of coordinates [27] and solved, yielding the set of tilt-corrected coordinates given by:

$$\begin{bmatrix} X_1 & \dots & X_n \\ Y_1 & \dots & Y_n \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X'_1 & \dots & X'_n \\ Y'_1 & \dots & Y'_n \end{bmatrix} \tag{13}$$

Solving for the coordinate matrix on the right-hand side yields:

$$\begin{bmatrix} X'_1 & \dots & X'_n \\ Y'_1 & \dots & Y'_n \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1} \begin{bmatrix} X_1 & \dots & X_n \\ Y_1 & \dots & Y_n \end{bmatrix} \tag{14}$$

In Equation (13) and (14), Y' and X' denote the tilt-corrected median location vector and the corresponding column number vector, respectively. A "smile model" is crucial during the profile cleaning procedure and computation of the asymmetry criterion. Although the curve formed by a human smile should resemble a hyperbolic cosine (if one approximates a lip as a spring), quadratic models form a good approximation at significantly lower computational expense. Thus a second order polynomial model of form $y = \hat{H}_1 x^2 + \hat{H}_2 x + C$ is fit to the rotated median data using the method of weighted linear least-squares as described in Equation (9-11). Note that two variables are modified in the aforementioned equations. Y' is substituted for Y and Ψ is instead given by:

$$\Psi = \begin{bmatrix} (X'_1)^2 & X'_1 & 1 \\ \vdots & \vdots & \vdots \\ (X'_n)^2 & X'_n & 1 \end{bmatrix} \tag{15}$$

Also note that the condition number [28] of the weighted Moore-Penrose pseudoinverse [29] exploited in the computation of the parameter estimate of the quadratic model is typically on the order of 10^{14}, suggesting that the fit is inherently likely to contain error due to high sensitivity. This provides further justification for the intensive care used to reduce the noise present in the data. The non-weighted Moore-Penrose pseudoinverse also has a comparable condition number, but its use is observed to result in a more biased fit. The mouth profile is then rotated counterclockwise by θ. The pixels in the rotated mouth profile that deviate more than 40 rows from the position predicted by the quadratic model are erased to reduce bias in the parameter estimate of the least-squares fit. This tilt correction and cleaning procedure is repeated until θ falls below an acceptable 0.1°.

After that, the asymmetry is quantified. Facial asymmetry can be gauged from assymetry in the the final quadratic model residuals. The asymmetry criterion proposed in this paper compares the mean squared error (MSE) of the quadratic model at both sides of its vertex (which is expected to be aligned with the midline of the face). This asymmetry criterion is given by the form:

$$\left| \frac{1}{R} \sum_{i=1}^{R} g(i) - \frac{1}{L-R+1} \sum_{i=R}^{L} g(i) \right| \tag{16}$$

$$where \quad g(i) = (Y'_i - f(X'_i))^2 \quad and \quad \frac{-\hat{H}_2}{2\hat{H}_1} \tag{17}$$

In Equation (16-17), Y', f, X', L, and $\hat{H}$ denote the final tilt-corrected median location vector, the final quadratic model, the tilt-corrected column number vector, the final median location vector size, and the parameter vector of the final quadratic model, respectively. Figure 5 demonstrates the processes described in this section for a normal and abnormal image of the second author.

Equation (16) yields a decimal value, which can be appended to a database during training or be compared to training values using Equation (4). The modified χ^2 test allows the algorithm to differentiate between facial paresis and innate facial asymmetry. As with the speech slur mini-test, the modified χ^2 statistics and raw asymmetry metric from this mini-test are stored in an internal SQLite database for future retrieval and use in comparison tests. The modified χ^2 test typically yields an output three orders of magnitude higher for abnormal (stroke) images compared to normal images, leading to a decisive separation between the two "classes." Yet, this test has good adaptability; comparison of the asymmetry values between two people with similar skin tones and the same gender gives quite reasonable J values, meaning that the

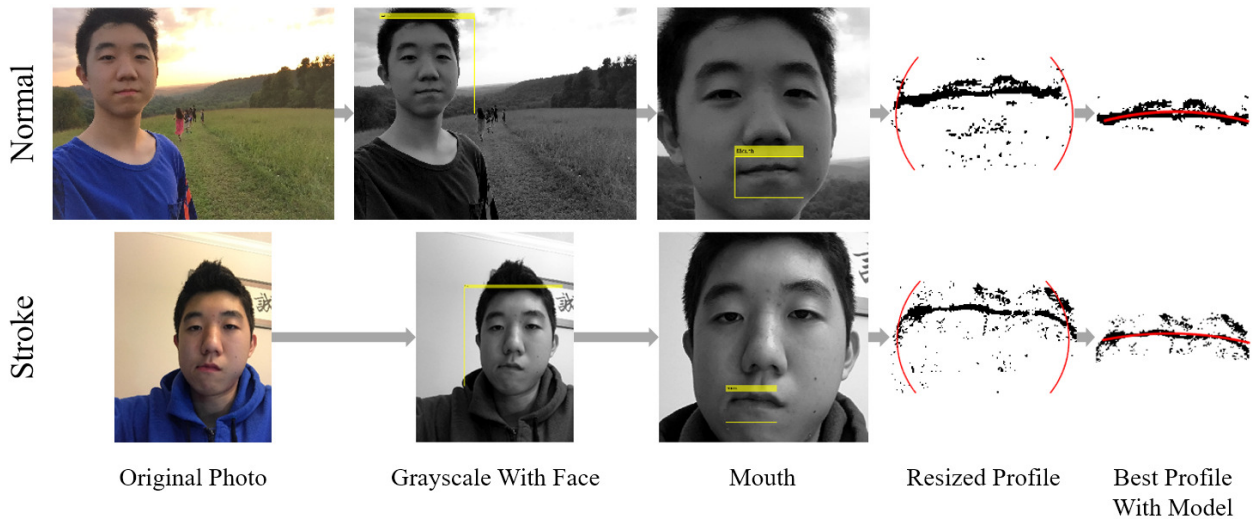

Fig. 5: Demonstration of Section III. D processes for a normal and abnormal image of the second author

facial asymmetry test may still be used for first time installers who suspect a stroke attack. In fact, the facial asymmetry test is not particularly sensitive to moderate aging either. Figure 6 demonstrates the flexibility of the processes described in this section in terms of aging and lighting for a normal and abnormal image of the first author.

E. Composite Susceptibility Metric

In order to develop a composite metric, a linear-in-parameter (LIP) model similar to that proposed in the face paresis test is considered. Such a model will have much lower computational expense compared to traditional nonlinear approaches such as neural networks and consume less memory than lookup table type approaches. It will also have the advantage of in-network training due to its mathematical simplicity and will be more rigorous than simply "counting the number of checked boxes." There are various stroke indicators that could be tested; however, in order to create a convenient yet effective self-examination, one must consider the simplest combination of variables that yields a model with high accuracy and acceptable false trigger rate. The following were the questions/variables under consideration:

1. Have you recently experienced blurred or tunnel vision?
2. Have you recently felt any numbness/tingling on one or both sides of your body?
3. Have you recently felt any weakness on one side of your body or face?
4. Have you recently felt SUDDEN confusion, dizziness, or nausea?
5. Speech slur?
6. Facial paresis?

Headache and vomiting could not be incorporated due to lack of relevant cross-tabulated data. Furthermore, they are common symptoms for a wide variety of illnesses and may bias the proposed model.

A "breadth-first" branch and bound approach [30] is used to solve this integer programming task. The aggregate LIP model could be written as:

$$\beta = \begin{cases} 1, & \zeta \geq K_1 \\ 0, & \zeta < K_1 \end{cases}; \quad \zeta = \sum_{i=1}^{6} \alpha_i Q_i \omega_i; \quad K_1 = 0.5 \tag{18}$$

$$where \quad \alpha_i \in [0,1],\ Q_i \in [0,1] \tag{19}$$

In Equation (18), Q_i, α, and ω denote the binary answer to question i, the binary vector corresponding to the current branch, and the weight vector, respectively. Note that LIP model outputs are continuous and are thus binarized through thresholding in order to construct a discrete classification. Equation (18) is fit to the data in the International Stroke Trial Database [31] in a way such that the weights are normalized and the impact of those who did not have strokes (only 2.2% of the dataset) on the model fit is equalized. To simplify model fitting, at each node of the branch and bound "search tree," only certain questions are used. The parameter vector is then estimated by solving a constrained, weighted linear least-squares problem of the following form:

$$\arg\min_{\omega} \frac{1}{2} \left\| W Q_{adj}^T \omega - W V^T \right\|^2 \ s.t. \begin{cases} \sum_{i=1}^{6} \omega_i = 1 \\ 0 \leq \omega \leq 1 \end{cases} \tag{20}$$

$$where \quad Q_{adj} = diag(\alpha_1, ..., \alpha_6) * Q; \tag{21}$$

$$W = diag(G_1, ..., G_6) \tag{22}$$

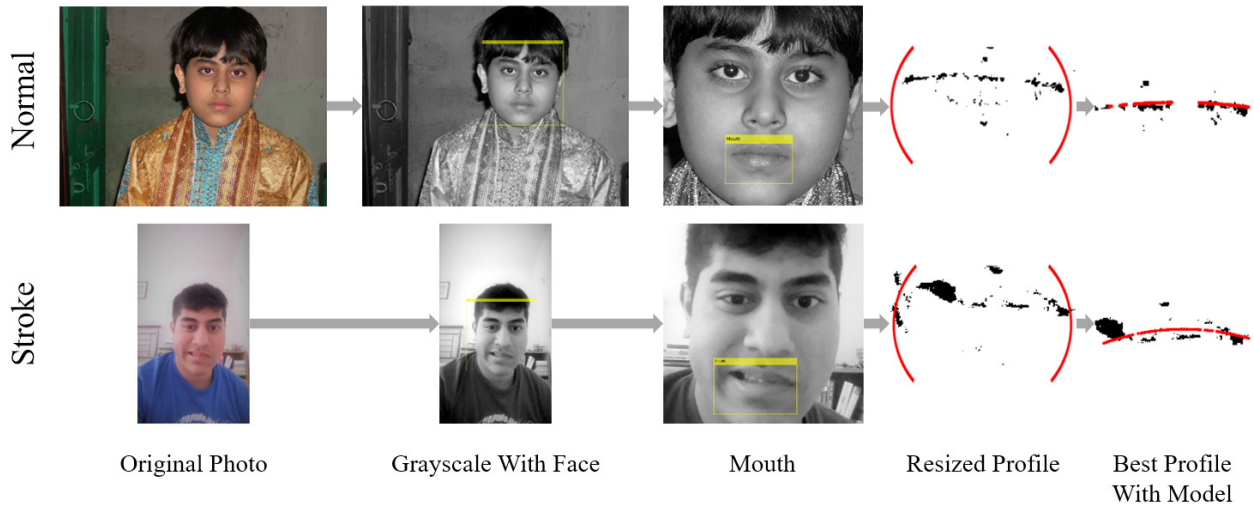

Fig. 6: Demonstration of Section III. D processes for a normal and abnormal image of the first author

$$and \quad G_i = \begin{cases} 1, & V_i = 1 \\ \sqrt{\frac{\sum_{i=1}^{n} V_i}{\sum_{i=1}^{n} 1 - V_i}}, & V_i = 0 \end{cases} \tag{23}$$

In Equation (20-23), ω and V denote the question weight vector and the binary truth case vector. Each nodal model is "trained" using 13.6 thousand data points. In order to determine the best model, the accuracy and false trigger rate of the various models are plotted on a Pareto objective tradeoff plot. Figure 7 exhibits such a plot for all of the possible nodes in the branch and bound "search tree." The selected model has a large Euclidian distance from the 50-50 tradeoff line (Figure 7) yet is compliant with the literature review. The final model selected for stroke susceptibility determination contains five of the six input terms and is given by:

$$\zeta = 0.35Q_1 + 0.31Q_3 + 0.05Q_4 + 0.09Q_5 + 0.21Q_6 \tag{24}$$

Equation (18) and (24) were validated on 5.8 thousand data points and can identify strokes 74.11% of the time but also false triggers 45.53% of the time. This alarmingly high false trigger rate is most likely due to the data being ill-representative of the population; all the persons in the survey in [31] had stroke-like symptoms and came to the hospital out of concern. This issue can be mitigated if a more inclusive survey was undertaken or if in-network training is pursued (see Section IV.B). The questions of relevance in Equation (24) are used at some point throughout the self-examination. Note that threshold K_1 can be used to fine-tune the accuracy and false trigger rate of the model. A sensitivity plot of the model with respect to K_1 is exhibited in Figure 8.

F. Testing Frequency Determination

To be effective, the self-examination should occur at properly timed intervals. That is, the time of the next self-test should be based off the time the user took the last one. In this paper this is done by recursively solving a stroke probability model modified from [32]. The original form of the probability model proposed in [32] as a function of time is given by:

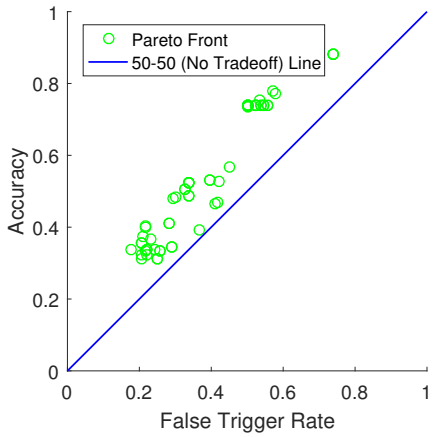

Fig. 7: Pareto front of all models

$$\rho(t) = 1 - (S(t))^{e^{(L_{com} - K_2)}} \tag{25}$$

Where (for males),

$$L_{com} = 0.0505L_{age} + 0.0140L_{sbp} + 0.3263L_{hyprx}$$
$$+0.3384L_{dm} + 0.5147L_{cigs} + 0.5195L_{cvd} \tag{26}$$
$$+0.6061L_{af} + 0.8415L_{lvh}$$

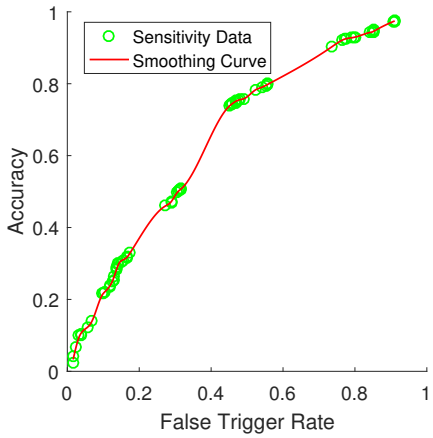

Fig. 8: Threshold sensitivity plot

or (for females),

$$L_{com} = 0.0657L_{age} + 0.0197L_{sbp} + 2.5432L_{hyprx}$$

$$-0.0134L_{sbp} * L_{hyprx} + 0.5442L_{dm} + 0.5294L_{cigs} \quad (27)$$

$$+0.4326L_{cvd} + 1.1497L_{af} + 0.8488L_{lvh}$$

In Equation (25), $S(t)$ and K_2 denote the probability of survival without a stroke for t years for an average individual and the value outputted by Equation (26, 27) if all inputs assume the average value for the given gender. In Equation (26, 27), L_{age}, L_{sbp}, L_{hyprx}, L_{dm}, L_{cigs}, L_{cvd}, L_{af}, and L_{lvh} denote the user's age in years, the user's systolic blood pressure in mmHg, the binary variable indicating the use of antihypertensive therapy, the binary variable indicating the presence of diabetes, the binary variable indicating the use of cigarettes, the binary variable indicating the previous diagnosis of any cerebrovascular disease, the binary variable indicating presence of atrial fibrillation, and the binary variable indicating the presence of left ventricular hypertrophy.

In [32], $S(t)$ was found from field data and only available in discrete instances of time. In order for Equation (25) and (26, 27) to be exploited for finding the testing frequency, a continuous functional form must be found for $S(t)$. Although the true model form is difficult to determine, $S(t)$ is approximated by a polynomial model. The order of the polynomial is found using Akaike Information Criterion corrected for small sample sizes (AICc) [33]. AIC is a widely used information theoretic criterion for finding the model order that yields the best tradeoff between goodness of fit and model complexity; it rewards goodness of fit while penalizing for increased number of model parameters [34]. AICc additionally corrects the AIC for sample size and is given by:

$$AICc = N\log(MSE) + 2\delta + \frac{2\delta(\delta+1)}{N-\delta-1} \quad (28)$$

In Equation (28), N, MSE, and δ denote the number of data points used to fit the model, the mean squared error

TABLE I: Goodness of fit of the polynomial models of $S(t)$

Gender	Model Form	MSE	R^2	AICc
Male	Linear	4.66e-5	0.949	-97.2
	Quadratic	2.21e-6	0.998	-124.5
	Cubic	1.60e-6	0.998	-123.4
Female	Linear	1.14e-5	0.971	-111.3
	Quadratic	9.75e-7	0.998	-132.7
	Cubic	4.04e-7	0.999	-137.2

TABLE II: Constants in Equation (29) for both genders

	A	B	K_2
Male	-5.326e-4	-4.509e-2	5.6770
Female	-2.583e-4	-3.939e-3	7.5766

of the fit, and the number of model parameters, respectively. First, second, and third order polynomial models are fit to the field data (different datasets for males and females) with the initial condition $S(0) = 1$. The goodness of fit for the various polynomials are shown in TABLE I.

AICc identifies the "best" model form by assigning it the most negative value. As shown in TABLE I, there is a close draw between the AICc values of the quadratic and cubic model for both genders. However, the quadratic model is favored in this paper due to its comparatively lower computational intensity and better "behavior." Substitution of the quadratic into Equation (25) yields:

$$\rho(t) = 1 - (At^2 + Bt + 1)^{e^{(L_{com} - K_2)}} \quad (29)$$

The values of constants A, B, and K_2 for both genders are given in TABLE II. Equation (29) is modified to yield the interval between consecutive tests. The method pursued in this paper is that of uniform probability spacing; self-examinations are requested after the probability of stroke increments a predetermined amount from the time of the previous self-examination. This interval can be written as the recursive solution to Equation (29) as given by:

$$t_n = \frac{-B + \sqrt{B^2 - 4A(1 - e^{\left(\frac{\ln(1-(\rho(t_{n-1})+T))}{e^{(L_{com}-K_2)}}\right)})}}{2A} \quad (30)$$

In Equation (30), T denotes the minimum probability increase that warrants a retest. This parameter is user-adjustable though $T = 1.5e-5$ seemed to give reasonable intervals. TABLE II exhibits some of the other constants used.

t_n needs to be converted into real time while taking into consideration the sleep cycle of the user for convenience. This is accomplished through use of modular arithmetic and "global clocks." The following equation accomplishes this task:

$$t_{c_n} = \begin{cases} t_b, & t_{u_n} \geq t_b \wedge t_s < t_b \\ t_w, & t_{u_n} < t_b \vee t_s > t_b \\ t_{u_n}, & else \end{cases} \quad (31)$$

$$where \quad t_{u_n} = (8766t_n + t_i) \mod 24 \quad (32)$$

$$and \quad t_s = (t_{n-1} + t_i) \mod 24 \tag{33}$$

In Equation (31-33), t_i, t_w, and t_b denote the installation time, the user's typical wakeup time, and the user's typical bedtime, respectively. Note that the aforementioned times are decimal values from 0-24 hours while t_n is measured in years. Users may voluntarily self-test at other times as well.

G. Emergency Contact

myCAST provides an expedient form of communication between potential stroke victims and emergency personnel. The innovative "Call 911 Now" button immediately dials emergency services for the user and simultaneously sends the user's GPS coordinates and any archived medical history (which is stored in the internal SQLite database) to the dispatcher. This information expedites communication between the user and EMTs in the event of a real emergency, reducing the delay between diagnosis and treatment. The emergency contact button is strategically placed on the home screen (Figure 1) and the self-examination results page (Figure 9) for ease of accessibility.

IV. Evaluation

A. Performance

The performance of the proposed application is tested on an Amazon Fire Phone v1.0. Note that myCAST is targeted at devices with a minimum SDK of 21 (but is backward compatible with SDKs as old as 17). In the short run, myCAST successfully appraises stroke susceptibility based on databases of test values. The proposed LIP model for the composite metric can identify strokes 74.11% of the time but false triggers 45.53% of the time. myCAST can also calculate and store data values for each of the mini-tests in an SQLite database, leading to a user-specific self-examination experience. An image of the "Recent Trends" page, which is generated from the archived data, is presented in Figure 10. myCAST can also send messages to and immediately dial to emergency services in the event of a detection. The app has low overall computational load, leading to good performance and expandability. Speech slur detection has slight issues when the user only marginally slurs, yet it is observed to work well otherwise. This is the same case with the facial paresis test.

B. Future Work

Stroke education is a cost-effective stroke prevention method and emergency response accelerator, making identification of risk factors and discussion of how to compensate crucial. In the future, stroke education should target education toward user risk factors and provide tips to mitigate their effects. A future study should correlate education media and information retentively. This would aid further development of the education section and improve quality of future informational mobile apps.

The algorithms presented in this paper are expandable to other applications. The speech slur test, for example, can be used to detect if the user is drunk. Then, based on the amount of speech slur present, the code could call a taxi with an algorithm similar to that of the intelligent contact button. However, these algorithms still require significant basic testing with a sizable test population to ensure reliability. In fact, the weights in the composite susceptibility metric still require some in-field adjustment, which could be done using a sequential estimation approach [35].

In addition to emergency contact, the mobile app should automatically send self-evaluation results to the user's physician. Tracking test results over extended periods could help physicians gain insight into occurrences of otherwise undetectable events like silent strokes (which may only show symptoms after extended periods). However, more rigorous testing is required to see if the false trigger rate of the self-examination is low enough to be trusted with automatic emergency contact.

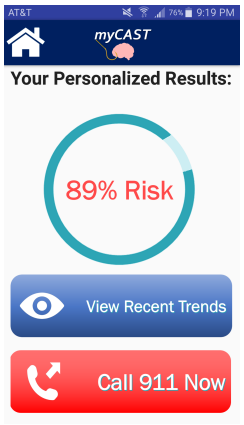

Fig. 9: Results tab

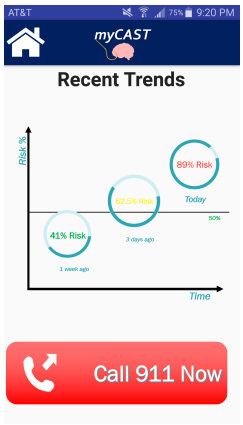

Fig. 10: Recent trends tab

V. Conclusion

In this paper, a novel mobile app, myCAST, was developed for combatting stroke fatalities. The concept was an all-encompassing, three-layer Android app for stroke education, self-examination, and emergency contact. The first layer was an educational section for stroke prevention designed to provide an interactive educational environment for teaching stroke symptoms and prevention. The proposed system would potentially enhance information retention and make symptoms more readily recognizable. The second layer performed automated self-examination of the user using a series of three mini-tests. The first mini-test was a questionnaire testing the user's perception of moderate stroke indicators. Slurred speech detection followed. The final mini-test detected facial paresis. Results were compiled into a single composite susceptibility score, which if exceeded a threshold, alerted the user to seek medical assistance. The final layer facilitated expedient communication through inclusion of an intelligent emergency contact button that sent the user's GPS location to dispatchers and medical history to EMTs.

It was found that the app was successful in the recognition of stroke symptoms, though its sensitivity and thus predictive ability were more limited. The LIP model for the composite metric could identify strokes 74.11% of the time but false triggered 45.53% of the time. myCAST could calculate and store data values for each of the mini-tests in an SQLite database, leading to user-specific self-examinations. This also allowed recent trends to be displayed and potentially analyzed; in the future, such data could be sent to doctors for further analysis. myCAST sent messages to and immediately dialed to emergency services in the event of a stroke detection. The app had low overall computational load, leading to good performance and a promising future of expandability.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to William Gibson for providing helpful discussions as well as a workspace for the project. The authors would also like to thank Dr. David Graham for helpful discussions and aid during the revision process. Additionally, the authors thank faculty and staff of Morgantown High School for their continual support.

REFERENCES

[1] The top 10 causes of death. (2017, January). Retrieved March 5, 2017, from http://www.who.int/mediacentre/factsheets/fs310/en/

[2] Stroke Facts. (2016, August 30). Retrieved November 16, 2016, from https://www.cdc.gov/stroke/facts.htm

[3] Ischemic Strokes (Clots). (2016, November 9). Retrieved March 05, 2017, from http://www.strokeassociation.org/STROKEORG/AboutStroke/TypesofStroke/IschemicClots/Ischemic-Strokes-Clots_UCM_310939_Article.jsp#.WN8qDjsrIeQ

[4] Transient Ischemic Attack (TIA). (n.d.). Retrieved March 06, 2017, from http://www.strokeassociation.org/STROKEORG/AboutStroke/TypesofStroke/TIA/Transient-Ischemic-Attack-TIA_UCM_492003_SubHomePage.jsp

[5] Stroke Treatment. (n.d.). Retrieved March 06, 2017, from http://www.strokeassociation.org/STROKEORG/AboutStroke/Treatment/Stroke-Treatment_UCM_492017_SubHomePage.jsp

[6] Hemorrhagic stroke. (2016, February 24). Retrieved March 06, 2017, from http://www.stroke.org/understand-stroke/what-stroke/hemorrhagic-stroke?gclid=CI6i-uG1gtMCFRtMDQodUBQLLQ

[7] Rathore, S. S., Hinn, A. R., Cooper, L. S., Tyroler, H. A., & Rosamond, W. D. (2002). Characterization of Incident Stroke Signs and Symptoms: Findings From the Atherosclerosis Risk in Communities Study. *Stroke,33*(11), 2718-2721. doi:10.1161/01.str.0000035286.87503.31

[8] Know your risk and prevent a stroke. (n.d.). Retrieved April 12, 2017, from https://www.strokeriskometer.com/

[9] Eichenbaum H, Yonelinas A.P, Ranganath C (2007) *The medial temporal lobe and recognition memory.*, Annu. Rev. Neurosci. 123-152

[10] M. Dzulkifli and M. Mustafar (2013) *The Influence of Colour on Memory Performance: A Review.*, Malays. J Med. Sci. 3-9

[11] Stroke. (n.d.). Retrieved April 22, 2017, from http://asha.org/public/speech/disorders/Stroke/

[12] Powell, Victor (n.d.). Princial Component Analysis explained visually. Retrieved April 13, 2017, from http://setosa.io/ev/principal-component-analysis/

[13] Lienhart R., Kuranov A., and V. Pisarevsky "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Object Detection.", Proceedings of the 25th DAGM Symposium on Pattern Recognition. Magdeburg, Germany, 2003.

[14] Levenshtein, Vladimir I. (February 1966). "Binary codes capable of correcting deletions, insertions, and reversals". *Soviet Physics Doklady*. 10 (8): 707710

[15] Pearson, Karl (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling"(PDF). Philosophical Magazine Series 5. 50 (302): 157175. doi:10.1080/14786440009463897

[16] SQLiteDatabase. (n.d.). Retrieved March 05, 2017, from https://developer.android.com/reference/android/database/sqlite/SQLiteDatabase.html

[17] Viola, Paul and Michael J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001. Volume: 1, pp.511518.

[18] Fmincon. (n.d.). Retrieved March 07, 2017, from https://www.mathworks.com/help/optim/ug/fmincon.html

[19] Jorge Nocedal and Stephen J. Wright (2006). Numerical Optimization. Springer. ISBN 0-387-30303-0

[20] Graythresh. (n.d.). Retrieved March 09, 2017, from https://www.mathworks.com/help/images/ref/adaptthresh.html

[21] Lloyd, Stuart P. "Least Squares Quantization in PCM." IEEE Transactions on Information Theory. Vol. 28, 1982, pp. 129137.

[22] Regionprops. (n.d.). Retrieved March 09, 2017, from https://www.mathworks.com/help/images/ref/regionprops.html

[23] Bourke, Paul (1997). "Polygons & Meshs".

[24] R. Keys, (1981). "Cubic convolution interpolation for digital image processing". IEEE Transactions on Acoustics, Speech, and Signal Processing. 29 (6): 11531160. doi:10.1109/TASSP.1981.1163711

[25] Statistical Analysis, Ya-lun Chou, Holt International, 1975, ISBN 0-03-089422-0, section 17.9

[26] Bühlmann, Peter; van de Geer, Sara (2011). Statistics for High-Dimensional Data: Methods, Theory and Applications. Springer. ISBN 9783642201929.

[27] Weisstein, Eric W. "Rotation Matrix." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/RotationMatrix.html

[28] Belsley, David A.; Kuh, Edwin; Welsch, Roy E. (1980). "The Condition Number". Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley & Sons. pp. 100104. ISBN 0-471-05856-4.

[29] Moore, E. H. (1920). "On the reciprocal of the general algebraic matrix". Bulletin of the American Mathematical Society. 26 (9): 394395

[30] Clausen, Jens (1999). Branch and Bound AlgorithmsPrinciples and Examples (PDF) (Technical report). University of Copenhagen.

[31] Sandercock, P. A., Niewada, M., & Czonkowska, A. (2011). The International Stroke Trial database. Trials, 12, 101. http://doi.org/10.1186/1745-6215-12-101

[32] Wolf, P. A., D'Agostino, R.B., Belanger, A.J., and Kannel, W. B. (1991). Probability of stroke: a risk profile from the Framingham Study. *American Heart Association Journals*

[33] Burnham, Anderson (2002). Model Selection and Multimodel Inference: A Practical Information Theoretic Approach. *Springer-Verlag*

[34] Akaike, H. (1974). A New Look at the Statistical Model Identification. *Springer Series in Statistics Selected Papers of Hirotugu Akaike*, 215-222

[35] Ferguson, Thomas S. (1967) *Mathematical statistics: A decision theoretic approach.*, Academic Press. ISBN 0-12-253750-5