

Available online at <http://www.mecspress.net/ijeme>

Personality Trait Identification Using Unconstrained Cursive and Mood Invariant Handwritten Text

Syeda Asra^a, Dr.Shubhangi D.C^{b,*}

^a*Appa Institute of Engineering & Technology, Visvesvaraya Technological University Belgaum, Brahmpur, Gulbarga 585102, India.*

^b*Visvevaraya Technological University Regional Centre, Rajapur Area, Kalaburagi 585104, India*

Abstract

Identification of Personality is a complex process. Personality traits are stable over time. Individual's behavior naturally varies from occasion to occasion. But there is a core consistency which defines the true nature. The paper addresses this issue of behavior. Graphology is normally a technique used to identify the traits. Accuracy of this technique depends on how skilled the analyst is. Although human intervention in handwriting analysis has been effective, but it is costly and prone to fatigue. An automation of handwritten text is proposed. Basically we have considered three important features in the direction of orientation of the lines : (i) up hill (ii) down hill (iii) constant line. Edge histogram and bounding boxes was used for feature extraction. Known classifiers like SVM & ANN are used for training and the results were compared. The results were about 98% for SVM & 70% with ANN. The analysis was done using single line.

Index Terms: Support Vector Machine; Artificial Neural Network; Personality Trait.

© 2015 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science.

1. Main Text

Human behavior[1] refers to the range of behaviors exhibited by humans and which are influenced by culture, attitudes, emotions, values, ethics, authority, rapport, hypnosis, persuasion, coercion and genetics. Human behavior is experienced throughout an individual's entire lifetime. It includes the way they act based on different factors such as genetics, social norms, core faith, and attitude. Behavior is impacted by certain traits each individual has. The traits vary from person to person and can produce different actions or behavior from each. Here introduce the paper, and put a nomenclature if necessary, in a box with the same font size as the rest of the paper. The paragraphs continue from here and are only separated by headings, subheadings, images and formulae. The section headings are arranged by numbers, bold and 10 pt. Here follows further instructions for

* Corresponding author. Tel.:0091-9986131522;
E-mail address: asrascholar@gmail.com

authors. person. Everyday conceptions of personality trait [2][10] make two key assumptions. First are stable over time. Most people accept that an individuals' behaviour naturally varies somewhat from occasion to occasion, but would maintain the core consistency which defines individual's true nature like the unchangeable spots of the leopard.

2. Personality Traits

Personality Traits are stable over time. Individual's behavior naturally varies somewhat from occasion to occasion but there is a core of consistency which defines the true nature. Traits directly influence behavior. There are several methods for identifying traits. Yet, trying to understand people's inner motivations is notoriously uncertain science. Personality identification offers us one of the few routes into this world. The Big Five personality traits are five broad domains or dimensions of personality [2] that are used to describe human personality. Contemporary research into the Big Five model [3] looks to assess how enduring these traits are over time, the degree to which there is a genetic component to these traits, and how these traits are correlated to life outcomes such as school performance. All individual differences can be summarized by five uncorrelated dimensions of personality.

2.1. Extraversion

This trait includes characteristics such as excitability, sociability, talkativeness, assertiveness and high amounts of emotional expressiveness.

2.2. Agreeableness

This personality dimension includes attributes such as trust, altruism, kindness, affection, and other pro-social behaviors. Common features of this dimension include high levels of thoughtfulness, with good impulse control and goal-directed behaviors. Those high in conscientiousness tend to be organized and mindful of details.

2.3. Neuroticism

Individuals high in this trait tend to experience emotional instability, anxiety, moodiness, irritability, and sadness.

2.4. Openness

This trait features characteristics such as imagination and insight, and those high in this trait also tend to have a broad range of interests. Correlation analysis is used to discover associations among measures. A strong association (high correlation) means people with one trait are likely to have second trait that the measures may overlap That is, they may, in part or wholly, be measuring the same trait or set of traits. All five factors show an influence from both heredity and environment. Openness shows 57% heritability while conscientiousness shows 49%, extraversion shows 54%, agreeableness shows 42%, and neuroticism shows 48%.

3. Genetic Influence on Human Psychological Traits

Among knowledgeable researchers, discussions regarding genetic influences on psychological traits are not about whether there is genetic influence, but rather about how much influence there is, and how genes work to shape the mind.

4. General Aspects of Image Processing[4]

4.1. Image Acquisition

This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves pre-processing, such as scaling etc.

4.2. Image Enhancement

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

4.3. Image Enhancement

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

4.4. Image Enhancement

Color image processing is an area that has been gaining its importance because of the significant increase in the use of digital images over the Internet. This may include color modeling and processing in a digital domain etc.

4.5. Wavelets and Multiresolution Processing

Wavelets are the foundation for representing images in various degrees of resolution. Images subdivision successively into smaller regions for data compression and for pyramidal representation.

4.6. Compression

Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data.

4.7. Morphological Processing

Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

4.8. Segmentation

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

4.9. Representation and Description

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

4.10. Object recognition

Recognition is the process that assigns a label, such as, “vehicle” to an object based on its descriptors.

4.11. Knowledge Base

Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted in seeking that information. The knowledge base also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications

5. Motivation

A study by American Management Association reveals that 39 percent of companies surveyed use personality testing as part of their hiring process [3]. Lawyers use personality testing for criminal behavior analysis, litigation profiling, witness examination and jury selection. Employers can reduce their turnover rate and prevent economic losses in the form of people prone to thievery, drug abuse, emotional disorders or violence in the workplace. Understanding Personality [5] what makes people tick becomes a key issue for anyone who works for others. This is particularly true, where we are trying to get the best out of people, allocate the right jobs to the right people and develop them. Software engineering abilities are desired skills for most employees [6]. Self-report inventories [7] contain transparent questions. While a significant amount of money is saved by crunching the interviewees down to the highest test scorers, the downside is that someone must score the tests. Even if the test is computerized, it takes time and money to create or choose the program that produces such a test. The test itself must be reviewed to determine what portions are most important to the company and available position. After applicants have completed the test, the employer must review the results to determine who to interview. An employer should carefully decide whether interviews or testing are the more costly way to evaluate potential employees, and weigh out the importance of profiling personalities via testing or face-to-face evaluations. Thus there is need to find an alternate method for personality identification.

6. Proposed Methodology

Graphology [8] is a study of handwriting (1871) and it helps in revealing the character and personality of the writer including his (or) her strength, weakness, and abilities. As handwriting comes from the unconscious, it contains a great deal of information, which can be useful for interpreting one’s character. ELANE describes Graph is Brain writing, the handwriting comes directly from the writer in a uniquely personal and individual way, irrespective of how the person has been taught to write; an expert graphology understand the styles of the different countries and languages and make allowances for — taught influences. Handwriting Analysis or Graphology [9] is a scientific method of identifying, evaluating and understanding personality through the strokes and patterns revealed by handwriting. Handwriting [10][11] reveals the true personality including

emotional outlay, fears, honesty, defences and many others. Professional handwriting examiners called graphologist often identify the writer with a piece of handwriting and prone to error and time consuming.

7. Fundamentals of Hand Writing Analysis[10]

- Basic Features of Handwriting
- Size and proportions
- General layout
- Direction of lines
- Degree of connection
- Regularity
- Rhythm (or evenness)
- Degree of broadness
- Speed of writing
- Form of letters
- Covering of space
- Shading
- Angle of writing (slope)
- Right and left tendencies
- Spacing
- Degree of attention
- Pressure
- Decoration/distortion

7.1. Is the Sample Valid

That is, has the style of the writing been affected by any external influences? E.g., an uneven writing surface; an awkward writing position; or written on the move (in a car or train, etc.). You cannot analyze a sample that is not reliable. For the purpose of checking slope and coverage it is more difficult to analyze samples that have been written on lined paper.

7.2. Size

There are many aspects to the size of writing and alone it doesn't indicate many things. Size (and especially the relative size of letter parts) needs to be considered along with other factors. There are some simple indicators however. Small writing is generally a strong indicator of a detailed, technical personality. Large rounded and dominant central case (see the later explanation of 'cases') letters indicate a friendly and sociable personality.

7.3. Letter-Word Slope

Is there a consistent slope to the letters and therefore the words in the sample? Check by drawing lines through the up and downward strokes. Backwards slopes indicate an introverted personality; forward slopes are extraverted. Upright strokes indicate a personality who is motivated by factors other than people, (ie neither extrovert nor introvert). The degree of slope reflects the degree of extraversion or introversion. The degree of consistency of the slope (ie parallel strokes) indicates the degree of emotional consistency.

7.4. Line Slope

Writing which rises to the right shows optimism and cheerfulness. Sagging to the right shows physical or mental weariness. (This applies to signatures sloping downwards also.)

7.5. Flow

One of the essential indicators, but like any other factor not to be used on its own. Generally, restraint is indicated every time the pen leaves the paper, and the converse applies. Gushing, eager, impulsive people have a more continuous flow of writing. Flowing writing has linked letters and sometimes linked words. Thoughtful, considered people will have fewer linked letters and no linked words. Artistic and conceptual people who like space and time around them will often have completely separated letters. (It follows that pressure at school on some children to 'write joined-up' - because the common view is that to do otherwise is 'not grownup' - is unnatural and often counter-productive. In any event continuity of flow does not correlate to intelligence.)

7.6. Spacing

Space between words indicates social attitude to others. Close words are a sign of sociability. Large spaces between words indicate the person is comfortable alone, and may even distrust others. Spacing between letters shows artistic spatially aware character, (artists, etc.)

7.7. Decoration and Distortion

Don't confuse the two. Decoration is intended; distorted is malformed and unintentional. Both are different to unfinished letters, which is a different indicator. Decoration is generally a sign that the writer wishes to be noticed more than he or she is at present. Malformed letters indicate a variety of things which must be dealt with individually. Unfinished words can be a sign of intelligence and impatience.

7.8. Intelligence and Impatience

The three cases divisions of the personality.

An essential aspect to analyze is the bias of the writing towards upper, central and lower cases. The upper case is the area in which the extended up-strokes are found in the tall letters like b, d, f, h, l, t, etc. The central or middle case is the central region occupied by letters with neither long up-strokes nor long down-strokes, such as a, c, e, m, n, o, etc. Which obviously means that the lower case is the area occupied by the extended down-strokes of letters such as g, j, p, q, etc?

Idealism, Religion Philosophy Upper. The central case contains most of the writing. The upper and lower cases are those which extend above and below the central body of each letter. The three cases represent the three aspects of our personality. If it helps you to assess the relative dominance of the three cases draw a horizontal set of 'tram-lines' through some lines of writing to mark the division between the three cases. Look at the relative dominance and extent of each of the cases. Upper our 'higher' selves, and thinking about religion and philosophy Central our mental and social approach to life Lower our physical aptitudes and attitudes. Look where the writing is mainly concentrated, and where the emphasis is; this is where the emphasis of the personality is too. Look for any encroachment from upper case to the line above or from lower case to the line below- it's a sure indication that the encroaching case is dominant. Also look at the central case to see if there is an upward or downward pull. An upward pull is best spotted if you see an arched pattern running through the central case. A downward pull is best spotted by seeing a pattern of troughs through the central case. The pull pattern in the central case also indicates the emphasis of the personality.

7.9. Attitude toward Environment

The position of the writing on a page within a margin frame shows how the potential employee regards his position or I space within the company framework.

7.10. Some Other Simple Indicators

Large broad upper loops are a sign of emotion. Right-pulled lower loops show interest for the good of others. Left-pulled upper case shows a fondness for the past. Uneven upper loops show changeable satisfaction, or disillusionment. Full left-pulled lower loops show physical fulfilment. Closed 'e's and other small closed central loops show secrecy. 'Stand-alone' or properly formed 's's at word ends show independence. Word-end 's's where top of letter is formed into downward right loop show a yielding or cooperative nature. Angular central case is a sign of an interest in ideas rather than people. Rounded central case shows interest in people. Uncompleted case letters, e.g. 'a' s, 'b's etc., (open when they would normally be joined loops) show a casual nature; very open shows propensity to gossip. Small writing is generally a sign of technical personality. Loops in angular writing show a potentially difficult character. 'T' cross strokes connecting a number of 't's with a single line shows speed of thought (but not a guarantee of correct thinking). Omitted 'i' dots and 't' cross strokes shows forgetfulness or carelessness. Position and style of 'i' dots show different things. Directly above, close to and in line means exacting. Ahead means active and thinking. Flicked shows sense of humour. Behind shows thoughtful. Inconsistent (varying positions) means a distracted mind.

8. Image Acquisition of Handwriting & Database Creation

Data samples of 500 in number were collected from people belonging to different works of life both equally from males & females. In order to make the results time and mood invariant the samples were collected in different days at different points of time. A4 paper with black ball point pen with a hard surface was used for writing. Three paragraphs were given to write. The images were scanned using laser jet scanner with a dimension of 2528x3507 pixels.

8.1. Image Handwriting Pre-Processing for Single Line

Since the samples collected were in the form of paragraph consisting multiple lines, hence there was a need to crop into single lines followed by de-noising[12] using salt and pepper method. All the images were De-noised & resized. A thresholding[13] function was used followed by gray to binary image conversion. Ten features[14] were extracted based on area perimeter, orientation, shape, maxima, minima points, upper left corner, lower left corner upper right corner and upper left corner of the bounding boxes were extracted. In order to find out the orientation of each line slope calculation was done.

8.2. Handwriting Segmentation

Bounding boxes[15] were used for finding the connected components.

8.3. Feature Extraction

Feature extraction is a technique of dimensionality reduction[16] from a high dimensional input datasets is to search for the projection of data on to smaller number of variables. This reduced output data is a transformation of the high dimension input data represented as feature vector.

8.4. Training Algorithm

- (i) Artificial neural network [16]

Delta Rule [Gradient Descent]

1. Initialize w = small random numbers
2. Until termination condition is met (error bound, or iterations of training examples)
4. Initialize all $\Delta w_i \leftarrow 0$
5. For each training example (x, t)
6. Compute the output of each node: $O(x)$
7. For each weight unit w_i : $\Delta w_i \leftarrow \Delta w_i + \eta (t - O(x))x_j$
8. For Each weight unit w_i
9. $w_i \leftarrow w_i + \Delta w_i$

8.5. Support Vector Machine

1. First, the input features are formulated as input vectors in some feature space.
2. Map these feature vectors to the higher dimension feature space using RBF kernel function.
3. Then a division global hyper plane is computed separate the feature space optimally to the classes of the training vector samples.

The network was trained using support vector machine. Out of 200 samples, 70% of it was used using for training and 30% for testing.

9. Co-relation Between Feature Detection and Personality

9.1. Up-Hill Line

Writing "uphill" reveals the applicant worthy of further Trait-match evaluation. He's optimistic, ambitious and cheerful. This is a good start, and, if the other traits bearup well under your scrutiny, a welcome addition to any office force or gang crew. Healthy mental energy & can stay busy active & constantly on the go.

9.2. Down –Hill Line

Writing "downhill" may be caused by a temporary depression, ill health, or physical fatigue. If this applicant has special skills necessary in your business, it would be wise to have him return on another day and obtain another sample of his writing. The first down slant sample could be just a temporary discouragement from job-hunting. If the second sample has the same down slanted appearance, it is a warning of an ingrained pessimism which keeps this applicant on the job-seeking circuit. Negative, disillusionment, constant disappointment.

9.3. Constant Line

The writer of a firm, straight, even baseline controls his moods, allowing him to go directly toward his goals without getting side-tracked. Reveals over control to compensate for an inner fear Healthy mental energy & can stay busy active & constantly on the go.

10. Results & Discussions

There are two phases of experiments. In the first phase we used 70% of the data set for training and 30% for testing in the second phase. 200 samples were collected equally from males and females. Samples were collected at different point of time from people belonging to different works of life hence making it mood[17] invariant analysis. Each writer was given an A4 paper and a black ball point pen to write three paragraphs.

These images were scanned with an HP scanner with 300 dpi. Since the image consisted multiple lines, hence we cropped it into single line as shown in Figure.1 and trained the network. Bounding boxes as shown in Figure.2 were used to determine upper left corner, lower left corner and centroid boxes region props. Regression line [18] was used to draw the line through these points for best fit. The network was trained using these features using SVM and ANN. In our work we not only worked on base line but also the line passing through the centroid of each bounding box and upper left corners of each bounding box to get accurate results.. As shown in Figure.3, Figure 4 and Figure 5 indicate orientation of handwriting in the uphill, downhill and line with a constant slope. The testing was done on a standard database and the results were 97%. The performance comparison with ANN and SVM classifier showed that SVM out performed ANN with 98% and ANN with 70%. Table 1 below shows the handwriting styles and its corresponding psychological personality traits[19][20] the writer may have. Table 2 indicate that the training time for SVM and ANN. There was a significant improvement in time when trained using SVM .Thus this system is ideal for online identification. A manual analysis was made about the personality trait[21] of the persons by continuously observing them for a span of six months under different situation and compared with the results, a strong correlation was found.

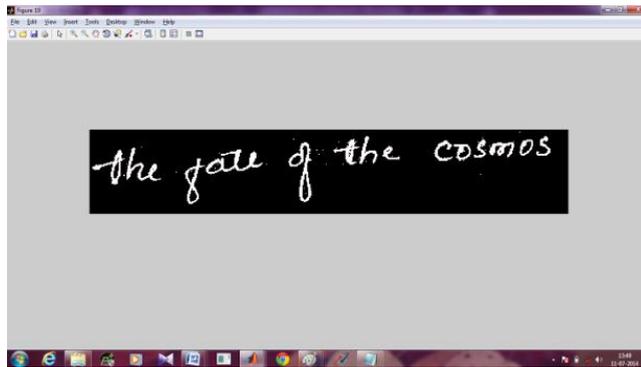


Fig. 1. Cropped Image of the Writing Sample

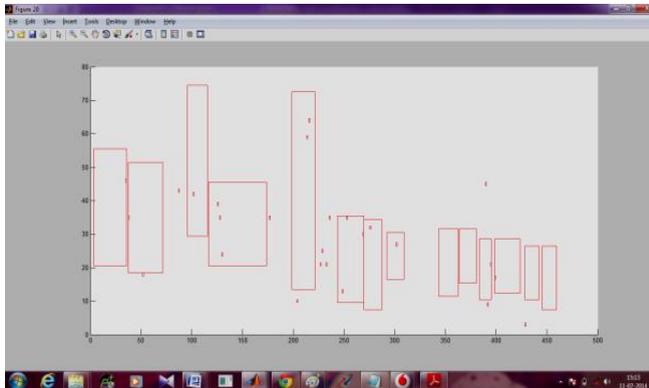


Fig. 2. Bounding Boxes for Connected Component

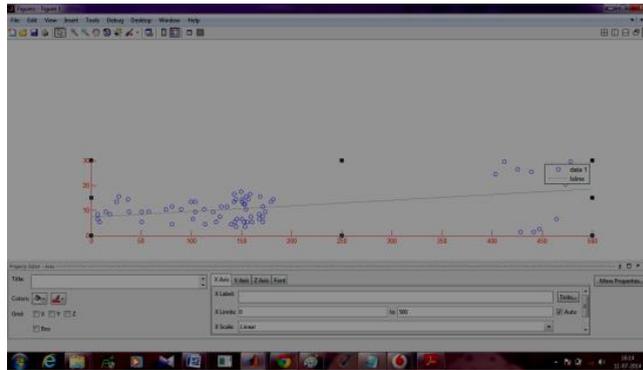


Fig. 3. Direction of Handwritten Text in the Up Hill

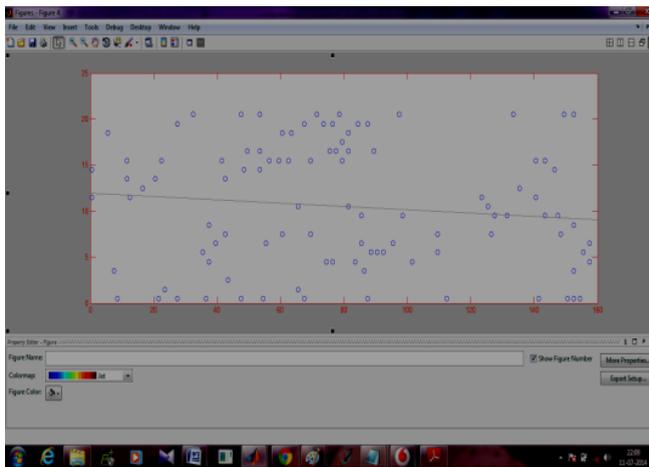


Fig. 4. Direction of Handwritten Text in the Downhill

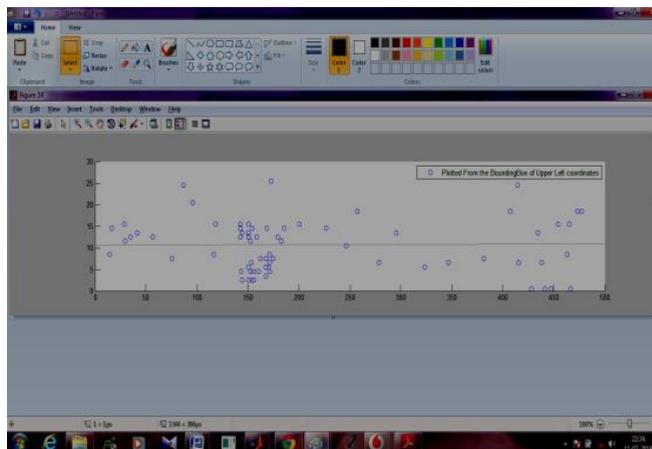


Fig. 5. Direction of Handwritten Text with the Constant Slope

Table 1. Personality Traits Predicted By Various Hand Writing Styles of Writer an Example of a Table

S.no	Writing Categories	Psychological & Personality Behaviour
1.	Baseline 	Healthy mental energy & can stay busy active & constantly on the go.
2.	Baseline 	Reveals over control to compensate for an inner fear Healthy mental energy & can stay busy active & constantly on the go
3.	Baseline 	Negative, disillusionment, constant disappointment.

Table 2. Average Time Cost of Three Retrieval Method

Algorithm	SVM	ANN
ime cost (min)	0.5	10.25

11. Conclusions

In this paper a novel regression line based orientation algorithm is used for finding the direction of orientation of lines .The proposed system is simple and effective and an alternate to manual detection of personality by a graphologists. The data set for testing was obtained from the standard data base and the results using SVM were 99%.

Acknowledgements

I would like to thank my parents my husband my kids and my guide for giving me encouragement in all kind of endeavours directly or indirectly. I further extend my thanks to the famous psychologist Gordon Allport.

References

- [1] Robert A. Rohm, Ph.D. | A powerful way to understand people an introduction of the DISC concept —Personality Insights, Inc. 800- 509-3472, 2010.
- [2] Gerrald Matthews, Ian J Deary, Martha Whiteman Personality traits | second edition 2003.
- [3] Susan J. Stabile | The use of personality tests as a hiring tool: Is the benefit worth the cost? || U.Pa. Journal of labor and employment law vol.4 2002.
- [4] Rafael Gonzalez | Digital image processing | 3rd Edition 2006.
- [5] Eric Garner — Understanding Personality Types: Managing people through their personality types | ISBN 978-87-403-00000 2012.
- [6] Timo Gnambs What Makes a Computer Wiz? Linking Personality Traits and Programming Aptitude

- Journal of Research in Personality, Volume 58, Pages 31-34 October 2015.
- [7] Jennifer Dodorico McDonald University of Cambridge| Measuring personality constructs the advantages and disadvantages of self- reports, informant reports and behavioral assessments| ENQUIRE Volume 1, Issue 1, June 2008.
 - [8] Beyerstein, Barry L. (Ed), Beyerstein, Dale F. (Ed) Amherst, NY, US —Evaluations of graphology, the study of handwriting analysis — Prometheus Books. (1992) 515 pp.
 - [9] Charles H. Trafford and Rand H. Nelson , Peterson Directed Handwriting |Handwriting A Complete Guide To Instruction Teaching Physical Patterns For Reading and Writing Fluency|ISBN 1-890666-3 5-1, 2003 .
 - [10] Shitala Prasad, Vivek Kumar Singh Akshay Sapre Handwriting analysis based on segmentation method for prediction of human personality using support vector machine, IJCA (0975 – 8887) Volume 8– No.12, October 2010.
 - [11] Syeda Asra, Dr. Shubhangi D.C —Personality Trait Identification –A Survey|IJCSN (2277-5420) Volume 3 ,Issue 2, April 2014 .
 - [12] Xinfeng Guo and Chunyan Meng Research on Support Vector Machine in Image De-noising Volume 8- No.02, pp.19-28, 2015.
 - [13] Ranjani J.J Bi-level Thresholding for Binarisation of Handwritten and Printed documents IEEE Conference Volume 9 Issue 1 2015.
 - [14] Shitala Prasad, Vivek Kumar Singh Akshay Sapre Handwriting analysis based on segmentation method for prediction of human personality using support vector machine, IJCA (0975 – 8887) Volume 8– No.12, October 2010 .
 - [15] P.Deepak and Smitha Suresh Design and Utilization of Bounding Box in Human Detection and Activity Identification Springer International Publishing Volume 338 ,pp 59-70, 2015.
 - [16] Yang Gao, Xuzhou et.al Dimensionality Reduction for Hyperspectral Data Based on Class-Aware Tensor Neighborhood Graph and Patch Alignment Volume 26 July 2015.
 - [17] Niriksha Nitika Jain et.al A Detailed Study on Artificial Neural Networks Discovery Volume 15 April 9 2014.
 - [18] Samuel, W. (1980) Mood and Personality correlates of IQ by race and sex of subject, Journal of Personality ND Social Psychology 38,993-1004. (2).
 - [19] Robert Nau Notes on Linear Regression Analysis Fuqua School of Business, Duke University 2014.
 - [20] Gerd Allport, Theory of traits| a critical review, 1999.
 - [21] Gerald Matthews, Ian J Deary, Martha Whiteman Personality traits| second edition 2003.

Authors' Profiles

Syeda Asra (born July 11, 1977) is an Indian researcher, keen in understanding the unique features within oneself and currently working as an Associate Professor in the department of Computer Science & Engineering in Appa Institute of Engineering & Technology, Kalaburagi, India.

Dr. Shubhangi DC is doctorate in Computer Science & Engineering .She has carried her research in Image Processing and Computer Vision. She is currently working as a Professor in the department of Computer Science & Engineering PG Centre Visveswaraya Technological University Kalaburagi, India.

How to cite this paper: Syeda Asra, Shubhangi D.C,"Personality Trait Identification Using Unconstrained Cursive and Mood Invariant Handwritten Text", International Journal of Education and Management Engineering(IJEME), Vol.5, No.5, pp.20-31, 2015.DOI: 10.5815/ijeme.2015.05.03