Matching Writers to Content Writing Tasks

Authors: Chandrashekhar Bhakuni, Khamir Purohit, Prabir Chakraborty, Ujjval Bhatt, Vikas Sardana (AMIL Nov 19 Group 3)

Problem Definition

We live in a world full of opportunities. Some of the present and future careers, especially the ones around digital marketing, content creation, and analytics did not exist even a decade ago. From content for business, to the business of content, there are exciting opportunities for content creators all around. In order to truly leverage these emerging vocations, and to convert them into lucrative careers, there needs to be a robust, scalable, and humanized system to identify the skills that specific content requirements warrant, and then find suitable writers for these roles.

For instance, when B2B companies engage writers for content marketing, they expect them to have subject matter knowledge. However, this search for qualified writers is not easy. Most companies rely on in-house resources, content marketplaces, advertising agencies, or freelance writers to outsource writing tasks. However, there is little they can do to find the best talent. This is because they are inundated with hundreds (if not thousands) of potential candidates with varying levels of technical knowledge, writing portfolios, and knowhow for the project at hand.

This project is an attempt to help businesses and communication agencies find the most suitable talent for niche writing jobs. At the same time, we will be helping experienced writers and Subject Matter Experts (SMEs) lend their services to content marketing projects, and receive a steady income. Such a nuanced approach will lead to top-quality content. The benefits are for all the stakeholders. The business gets the content they need, the content writer/ SME gets a chance to leverage his or her talent, while the reader gets authentic content that adds real value.

In this project, we propose to use computing power to rate write-ups created by academicians and experts with real insights gained from practical experience. This will not only benefit individuals, organizations, and agencies involved in content creation, but also benefit the World Wide Web by gradually improving the quality and relevance of information it holds.

We will use multiple data points around a writer’s ability, portfolio, academics, interests, work experience and more, and match them with available and upcoming content writing projects requirements, to arrive at the best match.

However, when one compares evaluation of written content with evaluation of images, we realise that great judgement is required for classification of write ups even by humans. Hence, building an ML model for analysing content would be a lot more challenging than one for image classification.

Literature Survey

The first step to evaluating a write-up (and hence its author) is to define what a "good" write-up is [1] . Writing for business domains is a techno-creative task. A writer needs to bring together the best of multiple worlds. First and foremost is the **subject matter understanding**. This helps her or him truly comprehend the business requirement, the ideas, the jargons, concepts, and processes.

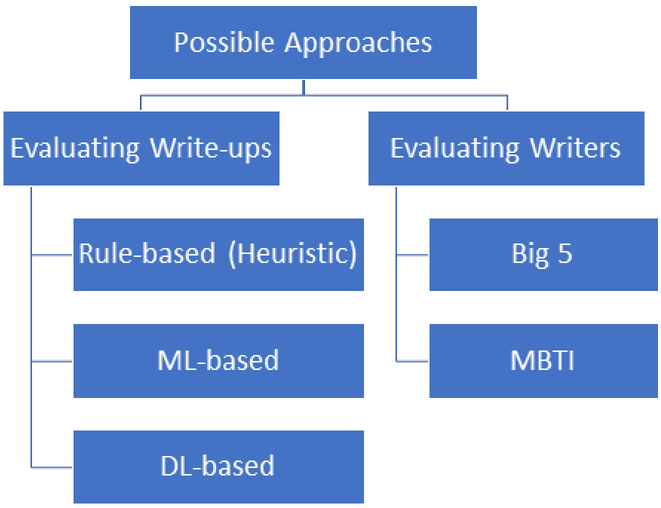
Secondly, the writer is required to have a **marketing bent of mind** to be able to subtly weave in a sales pitch or a call to action within the write-up. This warrants a thorough understanding of marketing communication and user behaviour, and the needs and expectations of the reader. In most cases of B2B content marketing, the target audience comprises a fairly evolved and technically sound audience that is looking for specific solutions. Hence, the writer should be able to communicate effectively and efficiently.

Lastly, the more traditional understanding of the term “a good writer”, is an ability to weave a **grammatically correct and structurally sound** narrative that follows all the commonly acceptable and expected norms of content formatting, referencing lexical, syntactical, fluency and prompt-specific features. An experience across content formats like blogs, white papers, technical case studies, web content, thought leadership articles, and social media platforms further helps the writer structure the content and leverage the specific features and benefits offered by that format.

In this study, we propose to explore multiple artificial intelligence and machine learning based approaches to meet the above-mentioned expectations. In order to meet our goal of identifying the best writer from a pool of thousands of potential writers, we believe that developing a simple and powerful system that operates at scale is a key outcome.

Possible Approaches

Research suggests that the current advances, both commercial and academic, are broadly based on two distinct approaches.



First approach - **Evaluate a Sample Write-up** to rate a writer. The criteria are:

Sr No.	Criteria	Explanation
1	**Grammar**	Sentence construction, punctuation, tenses, subject verb agreement, use of capitalization, etc.
2	**Tone of Voice**	Positioning of a business achieved through words. Technical, verbose, inwardly focused, informal, etc., are some examples
3	**Subject Matter**	Words, terminology, and references related to the subject being addressed - a critical aspect of the business content writing
4	**Formatting**	Consistency in formats for text, heading, bullet points, date, proper nouns, capitalization, etc., adds to the professionalism
5	**Structure**	Coherence of ideas and seamless flow of thoughts
6	**Call to Action**	If the write-up leads the reader to act. It is a good practice, not mandatory
7	**Product Placement**	Seamless pitch for a product/ service of a specific brand for which the piece is being written
8	**Depth and Recency**	Awareness of the right resources for the subject matter at hand. This adds to the relevance of the article for a B2B audience
9	**Ease of Readability**	How simple it is to read the given write-up. Vocabulary, length and nature of sentences determine this
10	**Redundancy**	How compact and to the point the writing is. A good piece will always be crisp. Devoid of long-winding and verbose sentences
11	**Hyperlinks**	Write-ups with hyperlinks to external references, reading material, sources, etc., are a good practice

Table 1: Technical Criteria to Evaluate a Write-up

Recent commercial applications use a personality-based approach to **Evaluate a Writer** by extracting personality traits from his or her writing:

Openness	**Conscientious**	**Extroversion**	**Agreeableness**	**Neuroticism**
Imaginative, Wide interest, Curious, Intelligent, Artistic, Unconventional	Organized, Disciplined, Planner, Goal oriented, not impulsive	Energetic, Forceful, Adventurous, Enthusiastic	Sympathetic, Straight forward, Compliance, Generous	Anxious, Tense, Worried, irritable, impulsive, shy

Table 2: Big Five Personality Traits

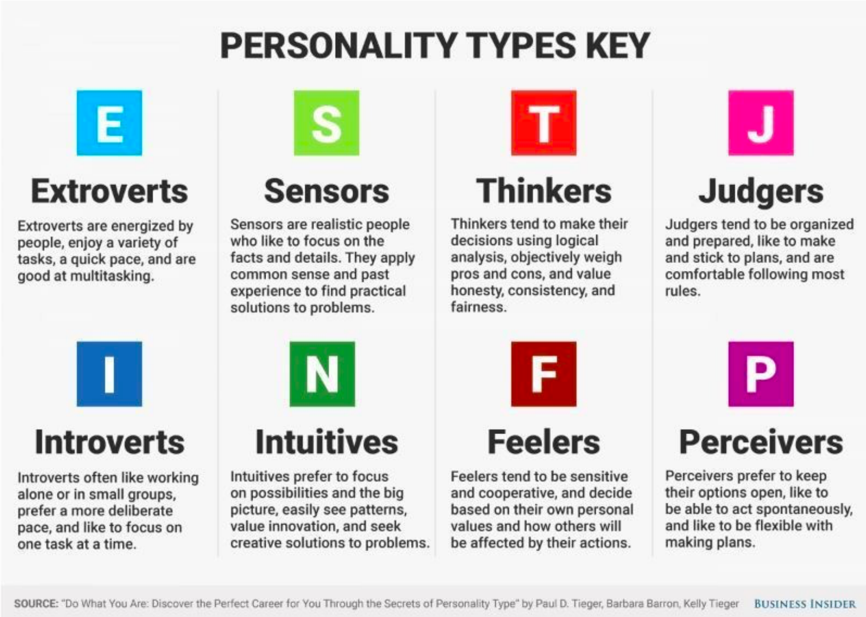


Table 3: Myers-Briggs Type Indicator

In both cases, the parameters thus extracted can be compared with the requirements of a given project (also arrived at by using the corresponding approach) and arrive at the best match.

Let us look at each of these in detail:

Evaluating Write-ups:

There can be multiple approaches to evaluate a write-up.

(a) **Heuristics** based or rule based approaches refer to the early attempts in building NLP systems. This required developers to have some expertise in the domain to formulate the rules and also required resources like dictionaries and thesauruses (example WordNet [2] , Open Mind Common Sense [3] , Regex (regular expression) based like StanfordCoreNLP's TokenRegex or Context Free Grammar (CFG) based JAPE (Java Annotation Pattern Engine) [4] or GATE (General Architecture for Text Engineering) [5] . Some of these are still used by NLP based products. This approach can help create the first versions of the solution which can be enhanced or used to create features for ML based solutions.

(b) **Machine Learning (ML**) based approach can be supervised, unsupervised or hybrid. They basically comprise three steps: Extracting features from text, Using the extracted feature representation to create a model, Evaluating and improving the model. Various ML concepts like Naive Bayes, Support Vector Machine (SVM), Hidden Markov Model (HMM), and Conditional Random Fields (CRF) etc., are used for solving the problems.

(c) **Deep Learning (DL)** based approach can help address shortcomings through Heuristic or ML based approaches. Recent advances in neural network based architectures like Recurrent Neural Networks (RNN) [6] [7] , Long Short Term Memory (LSTM) [8] , Gated Recurrent Units (GRU) [9] etc., helped solve many NLP problems. Recently Attention based architectures like Transformers [10] , BERT [11] and Autoencoders [12] are being employed.

The latest popular approach is Transfer Learning [13] , which is becoming popular with DL type of models. Here, a pre-trained neural network based model on a known task can be taken as a basis and fine-tuned for a new purpose-specific model. A different approach, exemplified in ELMo [14] , is feature-based training where a pre-trained neural network produces word embeddings which are then used as features in NLP models. In ULMFiT [15] , the authors demonstrated unsupervised pre training in NLP tasks by automatically generating labels from the data. They also showed that by fine tuning the pretrained model on just 100 labeled examples, they could achieve the same performance as a model trained from scratch on 10000 examples. OpenAI researchers, in their GPT paper [16] , showed how large but simple Transformer-like architecture can use the power of self-supervised pre training on large datasets and can then be adapted with minor tweaks to solve task specific problems.

One application of evaluating a write-up is used in Automated Essay Scoring (AES) systems. The AES systems rely on pre-determined features to evaluate the essays. The performance of such models is tied with the underlying features.

A 2011 Kaggle competition sponsored by Hewlett Foundation for Automated Essay Scoring [17] entailed the development of an automated scoring algorithm for student-written essays. The goal was to develop automated scoring systems to yield fast, effective and affordable solutions. The pre-graded essays were selected according to specific data characteristics. Finally, quality and reliability of participants' models were judged based on how well they could match expert human graders for each essay in terms of the scores. These scores were then checked against the manual score and Cohen's Kappa Score was generated for the result. Cohen's Kappa is a statistical measure used to evaluate the reliability of two raters who are rating the same quantity [18] . This score identifies how frequently the raters are in agreement. These AES systems assign a numeric score to the essay reflecting its quality, based on its content, grammar, and organization. These AES systems are, usually, based on regression methods. Typically, AES models were always prompt-specific. That is, models were required to be built specifically for each topic, using data from essays written to each of the particular prompts and scored by human raters. This process requires significant data collection and labelling of human reader scoring, which is a time-consuming and costly affair.

Some generic AES systems, which use limited features for evaluation and are not very prompt specific, are also being used. There are many such automated essay scoring systems [19] . E-rater [20] and Intelligent Essay Assessor [21] are two notable examples of AES systems. Automated Essay Scoring with E-rater® 2.0 used by Educational Testing Service (ETS) to evaluate the writing skills of students in their examinations such as GMAT, for example, uses 12 features including:

- Errors in grammar, usage, mechanics, and style (four features) – Model returns feedback about 33 errors across the four categories. Most of these features are identified using NLP techniques [22] . The features actually computed are the errors in the four categories and are calculated for each category by adding the total number of errors in that category and dividing by the total number of words in an essay.

- Organization and Development (two features): In addition to the various errors, the model identifies sentences in the essay that correspond to the following essay-discourse categories, using NLP: background, thesis, main ideas, supporting ideas, and conclusion [23] .
- Lexical Complexity (three features): Related to word-based characteristics:
  (a) The ratio of the number of word types (or forms) to tokens in an essay (referred to as type/ token) to count the number of unique words.
  (b) A measure of vocabulary level. Each word in the essay is assigned a vocabulary level value based on Breland's Standardized Frequency Index [24] and the fifth lowest standardized frequency index value is used to estimate the vocabulary level of the essay.
  (c) The average word length in characters across all words in the essay.
- Prompt-Specific Vocabulary Usage (two features): For each essay, six cosine correlations are computed between the vector of word weights for that essay and the word weight vectors for each score point. This feature indicates the score point level to which the essay text is most similar with regard to vocabulary usage. The second is the cosine correlation value between the essay vocabulary and the sample essays at the highest score point, usually 6 (referred to as cos. w/6). This feature indicates how similar the essay vocabulary is to the vocabulary of the best essays. Together these two features provide a measure of the level of prompt-specific vocabulary used in the essay.
- Essay Length (one feature): Essay length is the length of essay in terms of number of words.

The generic AES models with reduced feature-set that are fixed in nature (like the one explained in e-rater 2.0 above having 12 features) may not be prompt-specific and may be more uniform in nature across prompts. Such models, however, may not be able to give the nuances of content/ writing such as personality traits of a writer, etc.

Recently, various kinds of approaches including some based on NLP and machine learning models have been adopted for essay scoring. Most of these approaches consider essay rating as a classification, regression or preference ranking problem, where the loss function is the regression loss, classification loss and pairwise classification loss, respectively. Some other approaches [25] [26] propose a different method for e.g. a rank-based approach that utilizes a listwise learning to rank using Lexical features (statistics of word length, word level, unique words and spelling errors), Syntactical features (statistics of sentence length, subclauses, sentence level and mode, preposition, comma), Grammar and fluency features (word bigram and trigram and part of speech bigram and trigram) and Content and prompt-specific features (essay length, word vector similarity, mean cosine similarity of word vectors, and text coherence).

Some more developments based on commercially available NLP based tools such as Automatic Readability Tool for English (ARTE) [27] calculates readability formulas for texts. Another tool like Tool for Automatic Analysis of Cohesion (TAACO) [28] [29] calculates 150 indices of both local and global cohesion, including a number of type-token ratio indices (including specific parts of speech, lemmas, bigrams, trigrams and more), adjacent overlap indices (at both the sentence and paragraph level), and connectives indices. These approaches, thus, give numerous features for **evaluation of a sample write-up** but do not, necessarily, **evaluate a writer**. Moreover, coding and thresholds behind rating each feature is also not accessible for explaining the results.

Evaluate a Writer:

For the purpose of evaluation of a sample write-up, the approach in our project is similar to the aforementioned tools. In addition, our project aims to evaluate the writers by extracting personality traits from his or her writing.

One of the latest approaches commercially being employed is evaluation of the writer's personality. Rising popularity of this approach may be due to uncovering the hidden potential of writer's versatility and the range of quality can be displayed from use of words and combinations thereof.

The Big Five Personality Model is the most widely accepted and used model to assign a personal profile to an individual [30,31,32,33] . The Big Five Personality dimensions [34] are:

Personalities	**Scale**	**Significance**
Openness	Conservative and Traditional	Distinguishes imaginative, creative people from down-to-earth, conventional people
	Liberal and Artistic	
Conscientiousness	Organized and Hard Working	Control, regulate, and direct our impulses
	Impulsive and Spontaneous	
Extraversion	Contemplative	Enjoy humans, and seek excitement and stimulation
	Engaged	
Agreeableness	Competitive	Individual differences in concern with cooperation and social harmony
	Team working and Trusting	
Neuroticism	Laid back and Relaxed	Traits projecting the way people cope with, and respond to, life's demands
	Easily Stressed and Emotional	

Table 4: Big Five Personality Traits with Specifications

When dealing with the Big Five Personality Model, each individual can highly exhibit some of these traits together therefore meaning that the personality traits are not opposed to each other. A person can exhibit high symptoms of Agreeableness, Openness, while exhibiting little symptoms of Neuroticism [35]

Standardized tools are developed to measure an individual's personality based on the Big Five Personality Model. The tool is the NEO Personality Inventory, Revised (NEO-PI-R) which has 240 questions [36] . Since then, several studies were undertaken to infer personality based on the model. The most common measuring tool used is the Big Five Inventory (BFI), with 44 questions and NEO Five-Factor Inventory (NEO-FFI) with 50 questions [37] . However for our purpose getting the survey filled by authors may not be the most suitable way. So further studies were investigated to automatically predict personality.

Though the use of five personalities for writer's profiles is not researched more, a paper on Personality Recognition from Facebook activities [38] considered the approach by analyzing social media data shared by individual users. Facebook records a large amount of users' behavior expressed in various activities. This research was based on a dataset of 92,255 users who provided their Facebook likes and the results of their Big Five Personality test. The data includes 600 attributes LASSO algorithm was used to select the best features users' Big Five personality traits. The best accuracy level was reported for Openness and Extraversion, the lowest accuracy level is reported for Agreeableness while the accuracy levels of Conscientiousness and Neuroticism are in the middle.

An individual can exhibit high characteristics of more than one personality trait hence making it a multi-label learning task [39] . In machine learning, multi-label classification (MLC) is a form of classification problem but varies differently from other classification problems, in the sense that each sample can have several labels [40] . This varies from other classification problems that can have just one label and never two (i.e. an object can either be classified as dog or cat but never both) this is known as Multi-Class Classification. There are various real-world situations where MLC can be applied such as classifying a movie which can be both comedy and action genre .

The personality of an individual is stable through time and situation [41] , meaning the personality of an individual doesn't change online or offline, an individual that is sociable offline will be sociable online. However, as per profession writer the versatility is must to cater demand of the clients.

The other personality indicator is Myers–Briggs Type Indicator (MBTI) [42,43,44,45,46,47,48] , a forced-choice test based on Jung's personality typology which categorizes a person on four preferences:

- Introversion and Extraversion (attitudes): I's reflects before act, E's act before reflect.
- Intuition and Sensing (information-gathering): N's rely on abstract information, S's needs concrete information.
- Feeling and Thinking (decision-making): While F's decide with emotions, T's decided on logic and reasoning.
- Judging and Perceiving (lifestyle): J's prefer structure, P's like change.

A paper [49] on use of syntactic features to predict author personality from text used a new corpus, the Personae corpus, which consists of Dutch written language, while other studies focus on English. The technique can be transferable. The interesting thing was the data collection: Two ways data were collected (a) students had written the articles and (b) students had taken the MBTI test. The 200,000-word Personae corpus consists of 145 BA student essays of about 1,400 words were taken and MBTI tests for each of them were used to tag the author’s personality. To reduce the challenges, write up was about an opinion on a documentary on Artificial Life to keep genre, register, topic and age relatively constant. All students released the copyright of their text to the University of Antwerp and explicitly allowed the use of their text and personality profile for research, which makes it possible to distribute the corpus.

Sample Data

Content is all around us. From company websites and news portals, to endless social media feeds, personal blogs and editorials. However, all this data is unstructured, not labeled, and not tagged for machine learning. Moreover, content for business domains and topics is a niche.

In 2000, Reuters Ltd made available a large collection of Reuters News stories [50] for use in research and development of natural language processing, information retrieval, and machine learning systems. This dataset contains about 810,000 Reuters, English Language News stories from 1996-08-20 to 1997-08-19. The release date for this dataset is 2000-11-03 and it requires about 2.5 GB for storage of the uncompressed files.

This corpus, better known as "Reuters Corpus, Volume 1" or RCV1 is a well-known collection heavily used in the text classification community. This is highly relevant because we believe that one of our first objectives for an incoming reference text file will be to identify the “domain” it belongs to. Hence, we can use RCV1 for training on Text Classification as it carries business news articles. The stories in the Reuters Corpus are under the copyright of Reuters Ltd and/or Thomson Reuters.

The popular data corpus (original data as RCV1-v1) has been studied, enhanced, and cleaned [51] to generate a the corrected data as RCV1-v2 The corrected dataset covers the following categories CCAT (Corporate/Industrial), ECAT (Economics), GCAT (Government/Social), and MCAT (Markets). Further, Industry codes were assigned based on types of businesses discussed in the story. They were grouped in 10 sub-hierarchies, such as I2 (metals and minerals) and I5 (construction).

In addition, we will also be using MBTI Database [52] and Essay Rating Database [53] for further analysis.

About Proprietary Dataset

A proprietary untagged database with 5000+ write-ups is available for our project. This corpus is made accessible by a leading content writing agency in India.

The data is further categorized into various content formats like articles, white papers, web content, etc., with an average length of 800 words per write-up.

Given below are a couple of preliminary studies based on the proprietary data. The findings are presented in the form of an evaluation matrix.

Article Title	Grammar	Tone	Subject Matter	Formatting	Structure	CTA	Product Placement	Research	Readability	Redundancy Score	Hyperlinks	Class
The Importance of Life Insurance for Single Parents	8	8	8	8	7	0	NA	7	8	8	0	1
The next industrial revolution: powered by AI?	8	10	8	9	9	NA	8	7	9	8	9	1
Power Of Digitalization In Manufacturing	9	10	9	9	9	NA	9	8	9	9	10	1
Lust - The Desire Hypnosis	10	10	10	10	10	NA	8	10	10	10	0	1
Life Insurance and General Insurance: Explained	7	9	9	9	9	NA	8	8	7	7	5	1
Chhattisgarh Village On The Verge Of Being ODF	9	9	9	9	9	NA	10	10	8	9	0	1
Revolutionize The Way You Analyze Customer Experience Sentiment: A 5G Incight	7	5	8	10	8	8	8	8	3	8	8	0
What do customers look for when scouting for purchasing solutions?	5	3	3	3	2	8	7	5	2	5	5	0
Digital Agent/Agency Leader of the Year	5	6	7	8	7	NA	8	7	7	7	0	0
Mobile Banking	6	5	6	9	8	10	6	10	8	5	8	0
Online Lottery in India	6	4	7	3	5	5	7	8	4	4	4	0
Internet Banking	8	5	5	9	8	10	6	10	8	5	8	0

Table 5: Study 1 - Technical Evaluation of Sample Write-ups

For the 2nd study, we have taken 20 write ups written by three authors highlighted in by the same colour e.g., Rows in orange are write-ups by author 1, blue by author 2, and yellow by author 3.

-We can see that the algorithm returned identical personality traits for the write-ups written by the same author. Moreover, even after the write-ups were edited (shown as suffix 1) to satisfy the customer requirements, they still inherited the personality traits of the author.

Article	Score	Introversion	Extroversion	Intuiting	Sensing	Thinking	Feeling	Judging	Perceiving
Emkay Global	0	0.455	0.545	0.502	0.498	0.914	0.086	0.614	0.386
TCS iON	0	0.301	0.699	0.531	0.469	0.911	0.089	0.721	0.279
QSYS	0	0.422	0.578	0.394	0.606	0.915	0.085	0.761	0.239
SPJIMR	0	0.125	0.875	0.399	0.601	0.92	0.08	0.76	0.24
USS	0	0.383	0.617	0.333	0.667	0.903	0.097	0.691	0.309
Amazon Business	0	0.242	0.758	0.398	0.602	0.883	0.117	0.746	0.254
Bold	0	0.377	0.623	0.24	0.76	0.803	0.197	0.712	0.288
Godrej	0	0.463	0.537	0.191	0.809	0.738	0.262	0.579	0.421
ICICI Pru	0	0.233	0.767	0.656	0.344	0.903	0.097	0.764	0.236
V-Guard	0	0.361	0.639	0.145	0.855	0.808	0.192	0.56	0.44

Article	Score	Introversion	Extroversion	Intuiting	Sensing	Thinking	Feeling	Judging	Perceiving
Amazon Business	1	0.199	0.801	0.316	0.684	0.892	0.108	0.77	0.23
Bold	1	0.333	0.667	0.156	0.844	0.847	0.153	0.754	0.246
Emkay Global	1	0.419	0.581	0.648	0.352	0.926	0.074	0.631	0.369
Godrej	1	0.364	0.636	0.26	0.74	0.872	0.128	0.597	0.403
ICICI Pru	1	0.317	0.683	0.404	0.596	0.917	0.083	0.701	0.299
QSYS	1	0.473	0.527	0.367	0.633	0.904	0.096	0.665	0.335
SPJIMR	1	0.139	0.861	0.351	0.649	0.973	0.027	0.687	0.313
TCS iON	1	0.254	0.746	0.347	0.653	0.849	0.151	0.762	0.238
USS	1	0.308	0.692	0.496	0.504	0.877	0.123	0.712	0.288
V-Guard	1	0.424	0.576	0.164	0.836	0.776	0.224	0.63	0.37

Table 6: Study 2 - Technical Evaluation of Sample Write-ups

Our initial findings were that all Good writers need to be Thinkers (writing based on facts) and Judgers (prefer structures,rules, timelines). Introversion and Extroversion depends on the Domain or article requirement with Generally Extroversion linked with marketing, advertisement kind of article and introversion linked with scientific write-ups, technical papers.

Based on the preliminary study, we believe that our final approach would be a combination of the to sets of approaches. This hybrid is likely to provide us more comprehensive outputs.

Tentative List of Algorithms

- Multiclass classifications for rating quality of writing (for a given domain)
- Supervised Learning
- Cluster Recognition
- Linear Regression
- Recurrent Neural Networks (RNN)
- Long Short Term Memory (LSTM)
- Bidirectional Encoder Representations from Transformers (BERT)

References

[1] Weizhe Yuan, PengFei Liu, Graham Neubig. “Can We Automate Scientific Reviewing?: Page 3. Available at https://arxiv.org/pdf/2102.00176v1.pdf

[2] Miller, George A. “WordNet: A Lexical Database for English.” Communications of the ACM 38.11 (1995): 39–41.
Available at
https://www.researchgate.net/publication/31441225_Introduction_to_WordNet_An_On-line_Lexical_Database

[3] Catherine Havasi, Robert Speer, Kenneth Arnold, Henry Lieberman, Jason Alonso, and Jesse Moeller. “Open Mind Common Sense: Crowd-Sourcing for Common Sense” Common Sense Computing Initiative and MIT Media Lab
Available at https://www.aaai.org/ocs/index.php/WS/AAAIW10/paper/download/2011/2471

[4] Cunningham, Hamish & Maynard, Diana & Tablan, Valentin. (2000). JAPE: a Java Annotation Patterns Engine.
Available at
https://www.researchgate.net/publication/2495768_JAPE_a_Java_Annotation_Patterns_Engine

[5] Cunningham, Hamish & Humphreys, Kevin & Gaizauskas, Rob & Wilks, Yorick. (1997). GATE - a General Architecture for Text Engineering.. 29-30. 10.3115/974281.974299.
Available at
https://www.researchgate.net/publication/220708287_GATE_-_a_General_Architecture_for_Text_Engineering

[6] Williams, Ronald J.; Hinton, Geoffrey E.; Rumelhart, David E. (October 1986). "Learning representations by back-propagating errors". *Nature*. **323** (6088): 533–536. Bibcode:1986Natur.323..533R. doi:10.1038/323533a0. ISSN 1476-4687. S2CID 205001834.

[7] Karpathy, Andrej. “The Unreasonable Effectiveness of Recurrent Neural Networks”. May 21, 2015. Available at -
https://web.stanford.edu/class/cs379c/archive/2018/class_messages_listing/content/Artificial_Neural_Network_Technology_Tutorials/KarparthyUNREASONABLY-EFFECTIVE-RNN-15.pdf

[8] Ralf C. Staudemeyer, Eric Rothstein Morris (12 Sep 2019. “Understanding LSTM -- a tutorial into Long Short-Term Memory Recurrent Neural Networks”
Available at - arXiv:1909.09586v1

[9] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio. “Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling” Available at https://arxiv.org/abs/1412.3555

[10] Vaswani, Ashish & Shazeer, Noam & Parmar, Niki & Uszkoreit, Jakob & Jones, Llion & Gomez, Aidan & Kaiser, Lukasz & Polosukhin, Illia. (2017). Attention Is All You Need. Available at - arXiv:1706.03762v5

[11] Devlin, Jacob & Chang, Ming-Wei & Lee, Kenton & Toutanova, Kristina. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Available at - arXiv:1810.04805v2

[12] Dor Bank, Noam Koenigstein, Raja Giryes (12 March 2020). “Autoencoders” Available at - arXiv:2003.05991v1

[13] Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, Chunfang Liu (6 Aug 2018). “A Survey on Deep Transfer Learning” Available at - arXiv:1808.01974v1

[14] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, Luke Zettlemoyer (*15 Feb 2018 (v1), last revised 22 Mar 2018 (v2)*)”Deep contextualized word representations”
Available at - arXiv:1802.05365v2

[15] Jeremy Howard, Sebastian Ruder, (2018) “Universal Language Model Fine-tuning for Text Classification” Available at - https://arxiv.org/abs/1801.06146

[16] Alec Radford et al.,”Improving Language Understanding by Generative Pre-Training“ Available at - https://openai.com/blog/language-unsupervised/ and https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

[17] Kaggle competition The Hewlett Foundation: Automated Essay Scoring - Develop an automated scoring algorithm for student-written essays.[Online] available on https://www.kaggle.com/c/asap-aes

[18] Understanding Cohen’s Kappa Score With Hands-On Implementation. [Online] available at https://analyticsindiamag.com/understanding-cohens-kappa-score-with-hands-on-implementation/

[19] Handbook of Automated Essay Evaluation: Current Applications and New Directions. Routledge - Mark D. Shermis and Jill Burstein, editors 2013) [Online] available at https://www.routledge.com/Handbook-of-Automated-Essay-Evaluation-Current-Applications-and-New-Directions/Shermis-Burstein/p/book/9780415810968

[20] e-rater - Yigal Attali and Jill Burstein. 2004. Automated essay scoring with e-rater R v. 2.0. Technical report, Educational Testing Service. [Online] available at https://www.ets.org/Media/Research/pdf/RR-04-45.pdf)

[21] Peter W Foltz, Darrell Laham, and Thomas K Landauer. 1999. The Intelligent Essay Assessor: Applications to educational technology. Interactive Multimedia Electronic Journal

of Computer-Enhanced Learning. [Online] available at http://imej.wfu.edu/articles/1999/2/04/index.asp

[22] Burstein, J., Chodorow, M., & Leacock, C. (2003, August). CriterionSM: Online essay evaluation: An application for automated evaluation of student essays. In J. Riedl & R. Hill (Eds.). [Online] available at https://www.ets.org/Media/Research/pdf/erater_iaai03_burstein.pdf

[23] Burstein, J., Marcu, D., & Knight, K. (2003). Finding the WRITE stuff: Automatic identification of discourse structure in student essays. IEEE Intelligent Systems: Special Issue on Natural Language Processing, 18 (1): 32-39. [Online] available at https://www.researchgate.net/publication/3454073_Finding_the_WRITE_Stuff_Automatic_Identification_of_Discourse_Structure_in_Student_Essays

[24] Breland, H. M., Jones, R. J., & Jenkins, L. (1994). The College Board vocabulary study (College Board Rep. No. 94-4; ETS RR-94-26). New York: College Entrance Examination Board. [Online] available at https://onlinelibrary.wiley.com/doi/epdf/10.1002/j.2333-8504.1994.tb01599.x

[25] Kaveh Taghipour and Hwee Tou Ng - A Neural Approach to Automated Essay Scoring. [Online] available at https://www.aclweb.org/anthology/D16-1193.pdf

[26] Hongbo Chen and Ben He - Automated Essay Scoring by Maximizing Human-machine Agreement. [Online] available at https://www.aclweb.org/anthology/D13-1180.pdf

[27] Choi, J. S., & Crossley, S. A. *ARTE: Automatic Readability Tool for English*. NLP Tools for the Social Sciences. [Online] available at https://www.linguisticanalysistools.org/arte.html

[28] Crossley, S. A., Kyle, K., & Dascalu, M. The Tool for the Automatic Analysis of Cohesion 2.0: Integrating Semantic Similarity and Text Overlap. *Behavioral Research Methods*. [Online] available at https://link.springer.com/epdf/10.3758/s13428-018-1142-4?author_access_token=rMhwlSh0_7d8orvWZnaWBZAH0g46feNdnc402WrhzyqgLxPywVbXBKupRvVICdEKdhH5VXHqZIpT6gGW1KTfPcBQSwnCimj5EnAr57IhaQoCm70Wu6fr_iDRDA9cP66_jC0HJDtI_Q-FdVThn6q0TA%3D%3D

[29] Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods 48* (4), pp. 1227-1237. doi:10.3758/s13428-015-0651-7. [Online] available at https://www.linguisticanalysistools.org/uploads/1/3/9/3/13935189/10.3758_s13428-015-0651-7.pdf

[30] G. Farnadi, G. Sitaraman, S. Sushmita, F. Celli, M. Kosinski, D. Stillwell, S. Davalos, M. Moens and M. D. Cock, Computational personality recognition in social media, User Modeling and User-Adapted Interaction, vol.26, nos.2-3, pp.109-142, 2016. [Online] available at https://www.researchgate.net/publication/293194512_Computational_personality_recognition_in_social_media

[31] T. Ryan and S. Xenos, Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage, Computers in Human Behavior, vol.27, no.5, pp.1658-1664, 2011.[Online] available at https://isiarticles.com/bundles/Article/pre/pdf/32237.pdf

[32] C. J. Soto and J. J. Jackson, Five-factor model of personality, in Oxford Bibliographies in Psychology, D. S. Dunn (ed.), Oxford University Press, NY, 2013. [Online] available at https://www.researchgate.net/publication/264476432_Five-Factor_Model_of_Personality

[33] R. R. McCrae and P. T. Jr. Costa, The five-factor theory of personality, in Handbook of Personality: Theory and Research, O. P. John, R. W. Robins and L. A. Pervin (eds.), Guilford Press, NY, 2008. [Online] available at https://www.researchgate.net/publication/264476432_Five-Factor_Model_of_Personality

[34] Ateş, U. (2014). Inference of Personality Using Social Media Profiles, (June). [Online] available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.685.8347&rep=rep1&type=pdf

[35] Obinna Harrison Ejimogu (2017) Predicting Personality From Facebook Data: A Neural Network Approach, [Online] available on http://docs.neu.edu.tr/library/6688309210.pdf

[36] P. T. Costa and R. R. McCrae, Revised NEO Personality Inventory (NEO-PI-R) and NEO FiveFactor Inventory (NEO-FFI), 1992. [Online] available on https://www.researchgate.net/publication/285086638_The_revised_NEO_personality_inventory_NEO-PI-R

[37] O. P. John and S. Srivastava, The Big Five trait taxonomy: History, measurement, and theoretical perspectives, in Handbook of Personality: Theory and Research, Guilford Press, NY, 1999. [Online] available on https://darkwing.uoregon.edu/~sanjay/pubs/bigfive.pdf

[38] Asadzadeh Laleh; Rahimi Shahram (2017) Analyzing Facebook Activities for Personality Recognition, [Online] available on 10.1109/ICMLA.2017.00-29

[39] Obinna Harrison Ejimogu (2017) Predicting Personality From Facebook Data: A Neural Network Approach, [Online] available on http://docs.neu.edu.tr/library/6688309210.pdf

[40] Tsoumakas, G., and Ioannis, K. (2006). Multi-label classification: An overview. International Journal of Data Warehousing and Mining 3.3, 3 (3). [Online] available on https://www.researchgate.net/publication/273859036_Multi-Label_Classification_An_Overview

[41] Espinosa, M. J., and Rodríguez, L. F. G. (2004). Nuestra personalidad: En qué y por qué somos diferentes. Biblioteca Nueva. BIBLIOTECA NUEVA.[Online] available on https://dialnet.unirioja.es/servlet/libro?codigo=561377

[42] The 16 MBTI types by Isabel Briggs Myers

https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/the-16-mbti-types.htm?bhcp=1

[43] Gjurković, M., Šnajder, J., 2018. Reddit: A gold mine for personality prediction, in: Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media. pp. 87–97. https://www.researchgate.net/publication/325445581_Reddit_A_Gold_Mine_for_Personality_Prediction

[44] Heylighen, F., Dewaele, J.-M., 1999. Formality of language: definition, measurement and behavioural determinants. Interner Bericht, Center “Leo Apostel”, Vrije Universiteit Brüssel. http://pespmc1.vub.ac.be/Papers/Formality.pdf

[45] Mairesse, F., Walker, M., others, 2006. Words mark the nerds: Computational models of personality recognition through language, in: Proceedings of the Annual Meeting of the Cognitive Science Society. https://www.semanticscholar.org/paper/Words-Mark-the-Nerds%3A-Computational-Models-of-Mairesse-Walker/840a63afb5c0b428a6cfb45f6f4cb1555f205365

[46] Mairesse, F., Walker, M.A., Mehl, M.R., Moore, R.K., 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. Journal of artificial intelligence research 30, 457–500. https://www.researchgate.net/publication/220543010_Using_Linguistic_Cues_for_the_Automatic_Recognition_of_Personality_in_Conversation_and_Text

[47] Mehl, M.R., Gosling, S.D., Pennebaker, J.W., 2006. Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. Journal of personality and social psychology 90, 862. https://www.researchgate.net/publication/7046272_Personality_in_its_natural_habitat_Manifestations_and_implicit_folk_theories_of_personality_in_daily_life

[48] Mohammad, S.M., Turney, P.D., 2013. Nrc emotion lexicon. National Research Council, Canada. Tchokni, S.E., Séaghdha, D.O., Quercia, D., 2014. Emoticons and phrases: Status symbols in social media, in: Eighth International AAAI Conference on Weblogs and Social Media. https://researchswinger.org/publications/tchokni14.pdf

[49] Luyckx Kim, Daelemans Walter, Using syntactic features to predict author personality from text, Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), pp 146-149, 2008

[50] Reuters Corpora (RCV1, RCV2, TRC2) Available at https://trec.nist.gov/data/reuters/reuters.html

[51] Lewis, Yang, Rose, Li. “RCV1: A New Benchmark Collection for Text Categorization Research” Available at https://jmlr.csail.mit.edu/papers/volume5/lewis04a/lewis04a.pdf

[52] MBTI Database: Available online https://github.com/wiredtoserve/datascience/tree/master/PersonalityDetection

[53] Essay Rating Database:Available online https://www.kaggle.com/c/asap-aes/data