A Classifier to Predict Themes of Psychosis in English Literary Works

Abstract

Narratives of psychological and mental illness are exceptional because they challenge traditional reading practices which primarily focus on thematic analyses; sometimes, in the context of broader cultural concerns. Traditionally, such narratives are either read to validate their faithfulness to a real-life mental diagnosis, or to reflect on man's struggle with mental illness. However, it is nearly impossible to explore connections between literary texts that deal with psychosis unless of course these texts fall within the scope of a researcher's interest. What is lacking is a holistic and validated method that can be used to investigate representations of psychosis across large selections of literary works.

The researcher, utilizing data mining algorithms and the Weka software, developed a "classification model" based on a "training dataset" of quotes derived from Goodreads that dealt with all possible variations on the theme of mental disorder. The classifier identifies a set of attributes in the training dataset, then it was applied to a "test dataset" of 10 literary works to predict the occurrence of the previously identified attributes in these unclassified works. The present study outlines the process and results of implementing this classifier, and reflects on the potential of using text classification techniques in literary analysis.

Keywords

text classification, data mining, literary analysis, madness, psychosis, psychical disorders, thematics

نموذج تصنيف للتنبؤ بموضوعات الذهان في الأعمال الأدبية الإنجليزية

ملخص عربي

سرديات الذهان استثنائية لأنها تتحدى ممارسات القراءة التقليدية والتي تركز في المقام الأول على التحليلات الموضوعية، أحيانا، في سياق اهتمامات ثقافية أوسع من نطاق العمل الأدبي ذاته. وبشكل تقليدي، فإستقراء مثل هذه السرديات إما أن يكون بهدف التحقق من مطابقتها للواقع عند تصوير حالة عقلية، أو لمناقشة صراع الإنسان مع الإظطراب العقلي. ومع ذلك، فإنه من المستحيل تقريبا استكشاف الصلات بين النصوص الأدبية التي تناقش الذهان ما لم تكن هذه النصوص ضمن نطاق اهتمامات الباحث. ما ينقصنا إذن هو منهجية شمولية وذات مصداقية يمكن استخدامها للبحث في أساليب تصوير الذهان ويمكن تطبيقها على مجموعات كبيرة من الأعمال الأدبية.

لقد قامت الباحثة باستخدام خوارزميات التنقيب عن البيانات وبرنامج WEKA بتطوير "نموذج تصنيف" مبني على أساس "بيانات تدريبية" مستمدة من اقتباسات Goodreads والتي تناولت موضوع الاضطراب العقلي بصوره المختلفة. حددت خوارزمية التصنيف مجموعة من السمات في بيانات التدريب، ومن ثم قامت الباحثة بتطبيق هذه الخوارزمية على "بيانات الاختبار" والمكونة من 10 أعمال أدبية بهدف التنبؤ بظهور السمات التي سبق تحديدها في هذه الأعمال غير المصنفة. والدراسة الحالية تتناول طريقة تنفيذ الخوارزمية وكذلك نتائج التطبيق، مع اعتبار احتمالات استخدام أساليب تنقيب البيانات الكبيرة في التحليل الأدبي.

1 Introduction

1.1 Digital Humanities and the Rise of the Digital Literary Researcher

Research in the field of literary studies has been persistently informed, driven and highly motivated by a humanist perspective that predominately assumed some value or universal truth in literature which should be meaningfully investigated and interpreted through a close reading of literary texts (Love, 2010). Even when literary scholars have started to take interest in postmodernist and broader cultural theories as more flexible lenses through which they can approach and examine literary works, the emphasis was still on treating these texts as rich qualitative representations of human life that can never be subject to objective, deterministic or quantitative evaluations.

Unfortunately, this limited perspective of literary analysis and studies had led scholars in the field of literature to demonstrate a distrust of—bordering on hostility towards—digital applications that can be utilized in literary research. The major assumption is that the hypotheses and questions of literary research cannot be investigated using digital software, algorithms, or data mining techniques; that what Moretti (2000) designates "distant reading" actually undermines the significance and value of literary texts which a "close" or "cultural" reading can effectively highlight and communicate.

Not surprisingly though, current research in literary studies reveals a preoccupation with how literature is viewed and treated as "data" which can be processed, mined, and evaluated from within a "distant reading" framework and methodology; ultimately, and hopefully, contributing to traditional reading practices (Svensson, 2010; Clement, 2013; Flanders, 2005; Liu, 2012; Ramsay, 2003). Under the umbrella of what nowadays is labeled Digital Humanism, literary researchers have made use of several applications and techniques to enrich and broaden the impact of literary studies; thus paving the way for the same traditional questions about the value of literature to be studied using novel methods and approaches.

Forgoing the traditional literary scholar's interest in critiquing seminal works of literature while fundamentally grounded in a specific historical, social, or political milieu, the digital humanist engaged in literary studies adopts what can be designated a "blended approach," which proposes that literary reading should be "an examination of an aggregated ecosystem or 'economy' of texts" that will lead eventually to a "fuller understanding" of the cultural contexts of these texts (Jockers, 2013a). From the digital humanist's point of view,

> text analysis, visualizations, and data mining are just … tools, but they often provide the view the magnifying glass gives the user when he or she turns it upside down. These methodologies defamiliarize texts, making them unrecognizable in a way (putting them at a distance) that helps scholars identify features they might not otherwise have seen, make hypotheses, generate research questions, and figure out prevalent patterns and how to read them. (Clement, 2013)

Adopting such a framework then does not equate displacing a text from its literary or cultural contexts, rather augmenting some of its rarely unexplored dimensions, which contributes to a holistic reading; one that is not grounded in the humanist tradition alone.

One of these transformational methodologies that have found their way into the toolbox of the literary digital humanist is data mining. Data mining approaches, techniques and algorithms are used to translate either numerical (mostly structured) or textual (unstructured) data into knowledge and information that can be used to make decisions about a certain problem. As researchers in the field of literature, we are more engaged with the analysis of textual data which is a sub-topic under data mining, more appropriately labeled, text mining, in which one leverages various techniques to extract meaningful information from huge, mostly unstructured, datasets (Svensson, 2010; Aggarwal & Zhai, 2012; Hettinger, 2016). Of particular interest to the present researcher is text classification, a topic under text mining, which implements various machine learning algorithms to assign textual features to predefined categories or classes (Yu, 2008; Aggarwal, 2012).

Genre and influence studies, stylistic readings, gender-based critiques of literary texts, and theme analysis have greatly benefited from the integration of these digital methods and techniques into a larger, more comprehensive framework for literary studies, particularly those studies which practiced text classification to solve many literary questions. Among current projects in literary texts classification are ones which discussed literary style prediction. Cranenburgh and Koolen (2015), for example, used the SVM and regression algorithms to predict the literariness of 140 contemporary Dutch novels. They demonstrated that it is possible to automatically distinguish novels that are seen as highly literary from those that are seen as less literary, using simple textual features. Ciobanu et *al* (2013) implemented the forest tree classifier to predict the temporal space of historical Romanian Novels. Their dataset was reduced to ten main textual features on which they run the forest tree algorithm. Their results indicated that the classifier which they developed was able to predict whether or not a text belongs to the historical era it was written about. Another literary topic that has been investigated using text classification methods, is authorship attribution. Barufaldi *et al* (2009), to cite but an example, used the PPM-C classifier to predict the authorship of selected Brazilian literature utilizing predefined literary stylistic features. Similarly, Jockers (2013b) reported a 2008 study in which he, with a group of researchers, conducted an authorship analysis of the *Book of Mormon* employing the delta and nearest shrunken centroid (NSC) classification. The aim was to decide whether or not the book is a collaboration as well as validating certain extracts attribution to Smith, one of the contributors to the book.

1.2 Text Classification, Literary Themes, and Mental States

Basically, this paper maintains that digital methods; especially text-classification techniques and algorithms help literary scholars investigate, in addition to the topics explored in the previous section, thematic affinities among large collections of texts that span decades, sometimes centuries. Jockers and Mimno (2013), for instance, studied significant themes in large collections of nineteenth-century literature to uncover, through text mining and text classification techniques, how these themes can be mapped to other features like style, genre, and authorship. Some researchers have adopted topic modelling techniques to examine and evaluate literary themes across large content databases (Blevins, 2010).

The present researcher is mostly engaged with themes of madness, psychosis, and mental illness which are universal in the sense that men of letters since Shakespeare—and long before that—have explored in various ways. Literary narratives that toyed with the subject of madness, psychological or mental states are actually unique because they challenge traditional reading practices which depend on "close reading" to unmask the thematic concerns of the text; sometimes, in the context of broader cultural milieu since these literary accounts demand that our attention be directed towards the extraneous symptoms in total disregard to what it actually means to be mentally or psychologically disturbed. Close reading these narratives, then, might yield superficial outcomes that highlight the otherness of the individuals portrayed as suffering from mental illness. In addition, investigating the theme of mental or psychological illness is governed by the writing and reading traditions prevalent at a particular time. Thiher (1999), explains that "in literature as in life, we see that the mad inhabiting reality and the mad found in fictions live and experience their insanity in conformity with the explanatory paradigms that their era uses to understand madness."

Traditionally, such narratives are either read to validate their faithfulness to a real-life diagnosis, or to reflect on man's struggle with mental illness. Sometimes, gender, society, historical era, etc. are referenced as catalysts for a man of letters' engagement with these subjects (Harper, 1997; Jose, 2010; Donnelly, 2012). However, it is nearly impossible to fully explore connections between literary texts that deal with psychosis unless of course these texts fall within the scope of a researcher's interest. What is lacking is a holistic and validated method that can be used to investigate representations of mental illness across large selections of literary works, and while doing that uncover what it really means to be portrayed as mentally or psychologically afflicted: Does it entail being identified as an Other who, due to insanity or mental affliction, is presented as a spectacle to be observed? Or does it require that the author dive deep into such a mind to identify how it is similar to that of a sane person? At least, functioning in the same way, though the product might be different.

If narratives of the severely mental and psychological states are in themselves challenging to traditional reading, then maybe a machine learning, based on text-classification approach can help in uncovering hitherto undiscussed aspects of these themes. Hence, the main objectives of this paper are: 1) developing, implementing, and testing a classifier using training and testing sets which collectively consist of twenty one literary texts spanning mid-nineteenth, twentieth, and early twenty-first centuries, 2) outlining the process and results of implementing this classifier, and 3) reflecting on how these results can be used to contribute to a deeper understanding of psychological and mental states as depicted in literature.

The researcher focuses on three themes: entrapment, death, and self-reflection which are discussed in most of the research written about the selected fictional accounts. The characters suffering from mental and psychological problems in these narratives are reported to indulge in dark thoughts about being entrapped in a life which they can neither cope with nor understand, continuously contemplating taking their or other people's lives, and irrecoverably lost in their own minds playing and re-playing every emotion, response and encounter they experience (Höhn, 2007; Giang, 2014). In the following sections of this paper, the researcher will explain the text classification process and methods which she used to assign the literary dataset elements to each of the three themes (entrapment, death and self-reflection). The results of the classification process will then be demonstrated and discussed in the context of the overarching theme of mental and psychological illness.

2 Methods and Experimental Setup

2.1 Scope

The current study adopts a machine learning, text classification approach to determine whether or not the selected fictional narratives include references to three themes (entrapment, death, and self-reflection) which surface out as some of the well-studied and documented topics in the context of literary representations of mental illness and psychosis (Rieger, 1994; Donnelly, 2012). A total of twenty one fictional accounts were chosen for this experiment. Eleven of these narratives were transformed into, what in machine learning is called, a training dataset (seen and labeled data), while ten were treated as a testing set (unseen and unlabeled data).

While some of these literary accounts were classical or canonical in nature, others were popular bordering sometimes on confessional. From a genre perspective, they covered a broad spectrum (classic, gothic, young adult, psychological thriller, mystery, etc.). From a chronological point of view, these narratives span three centuries, from the middle of the nineteenth to early twenty-first century because the researcher wanted to draw conclusions about the occurrence of the themes without considering the cultural or historical aspects of the works. Table (1) below lists the fictional works used in the classification task presented in this paper:

#	Training Set	#	Testing Set
1	It's Kind of a Funny Story by Ned Vizzini	1	Tender is the Night by F. Scott Fitzgerald
2	Prozac Nation by Elizabeth Wurtzel	2	Shutter Island by Dennis Lehane
3	The Hours by Michael Cunningham	3	The Bluest Eye by Toni Morrison
4	Looking for Alaska by John Green	4	The Yellow Wallpaper by Charlotte Gilman
5	Fight Club by Chuck Palahniuk	5	The Quickening Maze by Adam Foulds
6	The Perks of Being a Wallflower by Stephen Chbosky	6	Mrs Dalloway by Virginia Woolf
7	One Flew Over the Cuckoo's Nest by Ken Kesey	7	Bell Jar by Sylvia Plath
8	I Never Promised You a Rose Garden by Joanne Greenburg	8	The Catcher in the Rye by J. D. Salinger
9	Girl Interrupted by Susanna Kaysen	9	The Tell-Tale Heart by Edgar Allan Poe
10	Sherlock Holmes by Arthur Conan Doyle	10	American Psycho by Bret Easton Ellis
11	The Icarus Girl by Helen Oyeyemi		

Table 1 List of fictional narratives used in this study

2.2 Text Classification Process

The text classification process implemented in this paper can be visually summarized in figure (1) below which shows how the database is fed with various formats of textual content that go through the four stages of preparation, pre-processing, classification and evaluation.

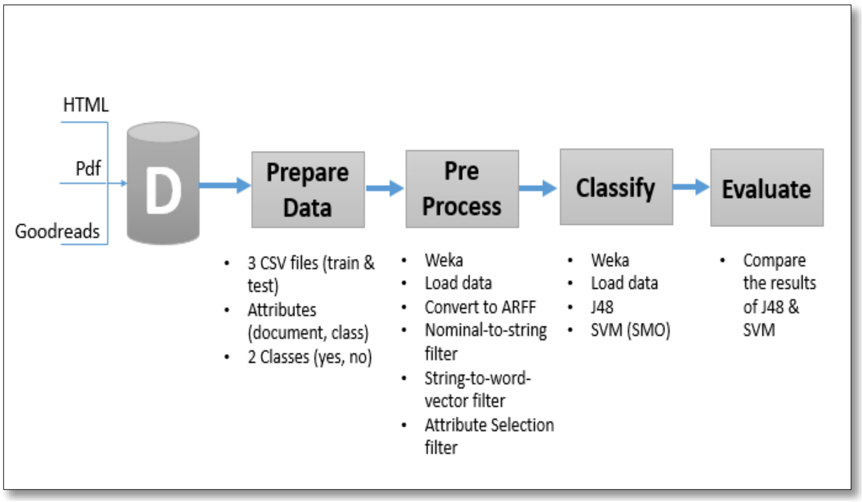

Figure 1 Text classification process

For this classification project, the researcher used Weka, which is an open source machine learning software developed by the university of Waikato in New Zealand. The software can be used to perform many data mining tasks including text classification either using its Java or explorer interfaces and allows for fast processing of data with access to different types of statistical and machine learning algorithms.

2.2.1 Data Preparation (Initial Pre-processing)

Quotes, from the eleven fictional works chosen to create the training set, were extracted from the Goodreads.com website using an online software (https://www.import.io). Basically, the online extractor software turned the quotes HTML pages into CSV files with multiple columns demonstrating detailed metadata for each quote. The researcher manually pre-processed the files to remove unwanted metadata as well as duplicated, foreign, and single-line quotations. In addition, she corrected typos, punctuation, and filled in ellipses using the original fictional works as reference. Then, all the extracted files were combined into one large, two-column CSV file that included 3522 quotes in one column with "text" as header, and "yes" or "no" labels in the second column with "class" as header. Because the researcher was developing a training set for three themes (entrapment, death, and self-reflection), she created three versions of this master file, one for each topic. The "yes" and "no" labels were input for the 3522 quotes accordingly. Each of the three files was opened using the Weka GUI and converted to an ARFF format, which is needed to perform classification tasks planned for this project.

As for the ten fictional works chosen for the testing set, the researcher converted the HTML pages and PDF files, which were the digitized formats in which these works were found, into one large two-column CSV file using a host of online extractors, and Excel formulas. Each row of the CSV file contained an extract that is no longer than 20 lines under the header "text". The second column's header was named "class" and "?" was input as label so that when processed in Weka, the missing values will be replaced with predictions (either "yes" or "no"). The final cleaned version of the testing set included 1621 extracts. Because the researcher is predicting the occurrence of three topics (entrapment, death, and self-reflection) in the testing set, three versions of this file were created, converted to ARFF, and merged with the corresponding training file. The final product for this manual pre-processing stage are three ARFF files containing the same training and testing sets, but the training extracts are coded differently in each file in accordance with the theme being examined. For the sake of simplification and clarity the files will be henceforward referred to as (dataset A for Entrapment, dataset B for Death, and dataset C for Self-reflection)

2.2.2 Pre-processing in Weka

Literary data pre- processing in Weka is then performed to help develop a classifier which can learn the information essential in the allocation of literary extracts to one category (theme) or another. It is very important to note that the three files containing the literary datasets, in their original state, are represented as nominal data which classification algorithms cannot read, hence cannot learn properly. This data had to be converted into a machine-readable format. Therefore, the merged files created in the previous step were imported in Weka, one by one, and transformed into word vectors using first a "NominalToString" filter, then a "String-To-Word-Vector" filter, the latter, as figure (2) demonstrates, performs a predefined process which entails 1) splitting the text strings into meaningful elements like words or N-grams, which is called Tokenization; 2) lowering the uppercases of text strings to avoid listing the same word as two because of the typographical mark; and 3) removing stop words (the, is, are, etc.) from the literary extracts due to their high occurrence that might impact how the classifier learns the information. The result of this pre-processing step is a list of words that occur in all of the literary extracts, and with enforcing the TF-IDF (term frequency–inverse document frequency) option in Weka, the list would not just show the raw frequencies of words across all the dataset, but it will demonstrate statistically how important these words in particular documents compared to others.

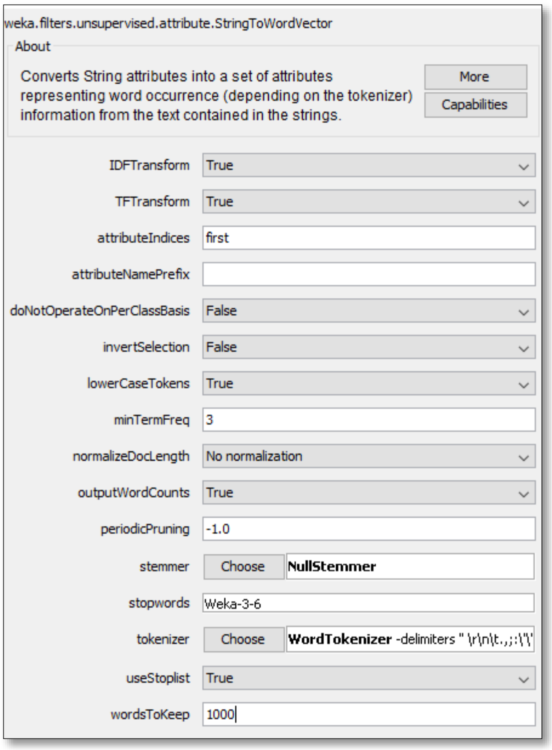

Figure 2 String-to-word-vector options

Definitely, this helps in the following feature extraction step, where the list of words (now identified as textual attributes) is filtered using the "AttributeSelection" filter in Weka. What this filter does is that it ranks the words in the list and chooses the most important ones for the classification process. This reduces the number of attributes (words) and helps the classifier learns faster. At this stage, the three datasets are ready for classification.

2.2.3 Developing and Evaluating the Classifier

This is the main step in the text classification process proposed in this paper. The researcher tested two implementations of two popular algorithms, J48 (for the C4.5 algorithm) and SMO (for the Support Vector Machine (SVM) algorithm). J48 simply generates decision trees from the training data through selecting the attribute/s which successfully divide its instances into subsets that can be found in one class more than the other. Basically, it splits the data into nodes according to a set criteria and creates a visual network of the dataset. The SVM does almost the same thing; i.e. class prediction, but with much precision. It maps the attributes as belonging to one class or another by creating a very clear distance (gap) between the two classes. In this experiment, the researcher ran the algorithms on the literary dataset using three validation methods: using training set, percentage split and 10-fold cross validation, which, as the following section will demonstrate, yielded different results.

3 Results and Discussion

The results of pre-processing the three datasets (A, B, and C) in Weka is a reduced number of attributes (words), especially after the "AttributeSelection" filter was applied. Table (2) below shows the number of the remaining attributes in each dataset:

Dataset	*Theme*	*NO. of Attributes before applying Filter*	*NO. of Attributes After applying Filter*
A	Entrapment	1076	31
B	Death	1892	222
C	Self-Reflection	1001	117

Table 2 Remaining attributes in datasets, A, B, and C

It is these reduced datasets that were used in the classification process. Actually, data reduction was helpful as it allowed the researcher to examine the resultant word list and visualize it. For example, when inspecting a visual representation of dataset B, the researcher notes how the sum of words like “dead”, “die”, and “kill” across the dataset was the highest compared to other words denoting death. Figure (3) is only a simple visualization of a subset of dataset B.

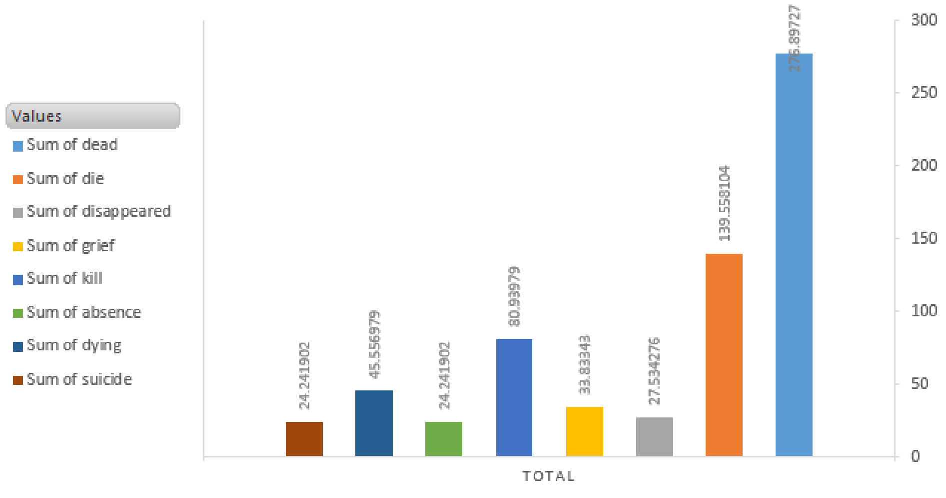

Figure 3 Visual representation of a subset of attributes in dataset B

However, if we stop here and accept the frequency of the words in the reduced word list as an indicator of the fictional characters’ obsession with death, for example, then we are glossing over the significance of the words in their contexts and most importantly their even distribution among the extracts making the dataset. Does the high frequency of a particular word mean that the dataset deals with the theme or idea associated with the word? Besides, and as figure (4) shows, the words might have nothing to do with the topic or theme being studied. The researcher in dataset (A) is exploring how the characters suffering from psychological or mental problems talk about how they feel entrapped in their world and lives, however, the reduced world list does not contain terms that denote entrapment.

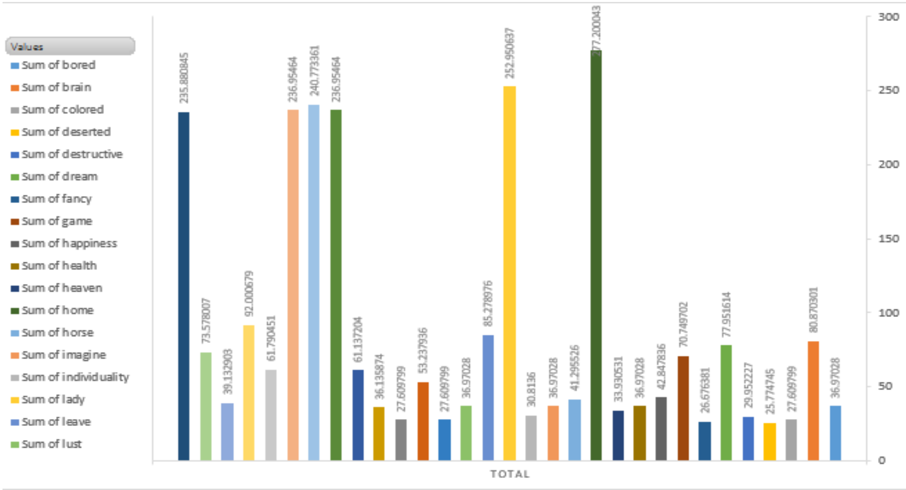

Figure 4 Visual representation of a subset of attributes in dataset A

The frequency of words should not be taken as the only indication of the presence of a theme in a particular literary text. One has to mine the texts for other clues that might be useful in deciding whether a certain literary extract actually deals with a particular theme. Most importantly, whether this theme can be later identified as a major theme in these texts.

Running the J48 and SVM classifiers on the three reduced datasets (A, B, and C) described previously, while using the training set as validation, resulted in predictions listed in table (3) below:

Theme	*Algorithm*				*Actual*	
	J48		**SVM**			
	Predicted		**Predicted**			
	Yes	**No**	**Yes**	**No**	**Yes**	**No**
Entrapment	44%	56%	30%	70%	27%	73%
Death	34%	66%	22%	78%	17%	83%
Self-reflection	63%	37%	71%	29%	76%	24%

Table 3 J48 & SVM classification and output predictions

To a casual observer, the results in the above table might seem of little value statistically speaking. One notes how the values of the algorithmic predictions differ from those representing the actual, manually coded, classifications. But, we are not looking for identical figures. What one might elicit out of this is that; firstly, algorithms were able to learn the data and approximate the values of the manual classification to a certain extent especially with SVM; and secondly, one might start to draw inferences about the three themes and how self-reflection, for instance, seems to be frequent in dataset C, where the “yes” label scored 63% with J48, 71% with SVM, and 76% in actual classification. Moreover, reading the values of the “no” label for each theme should motivate us to take into consideration how the negative values can be viewed as meaningful as the positive ones.

The output of the classification process is not just limited to the label values reported above. Tables (4) demonstrates the accuracy and error percentages of the algorithms’ implementations as extracted from the Weka software, using a 10-fold cross validation method of evaluation. Basically, this validation method checks the performance of the algorithms and provides us with a more accurate estimate than the other techniques. This method splits

> the dataset into k-partitions or folds [k = a value]. Train a model on all of the partitions except one that is held out as the test set, then repeat this process creating k-different models and give each fold a chance of being held out as the test set. Then calculate the average performance of all k models. (Brownlee, 2016)

Theme	***Algorithm***			
	J48		**SVM**	
	Accuracy	**Error**	**Accuracy**	**Error**
Entrapment	81.2992 %	18.7008 %	81.6929 %	18.3071 %
Death	85.0785 %	14.9215 %	92.4084 %	7.5916 %
Self-Reflection	70.8401 %	29.1599%	90.013 %	9.987%

Table 4 Results of running J48 and SVM using the Training Set

There is a slight difference between the accuracy values for the J48 and SVM when run on dataset A (entrapment), which indicates that the performance of both algorithms is moderately high on this dataset. Both algorithms succeeded in classifying the extracts as either containing a reference to the theme, entrapment, or not. The SVM algorithm performed better with datasets B (death) and C (self-reflection) scoring an accuracy of 92% and 90% respectively.

How can we interpret these scores in the context of literary studies? Actually, such values seem to indicate that there are close affinities between the texts chosen for this research, despite the fact they belong to different sub-genres, and eras. If the dataset features were diverse, the classifier would not have been able to predict with such precision. Then, there must be an underlying link between narratives portraying mental and psychological issues, and such link can be observed and studied if it is visualized. Text mining techniques assist in visualizing these hidden links so that the literary researcher can easily incorporate them into his interpretive or critical process.

What this exercise tells us, essentially, is that the thematic features identified in the chosen literary texts, when examined as a whole and from a distance, turned into a statistical pattern, which differs from the statistical patterns existing in other literary works that do not reflect on the themes of mental disorder. Error values become as significant as accuracy values, because, for a literary researcher, these open the door for an interpretive engagement with the text. For example, we can ask ourselves why did the classifier failed to classify any of the extracts taken from Bret Easton Ellis's *American Psycho* (1964) as self-reflective although the novel is written almost mainly in the first person. It might have been due to the way the main character of *American Psycho* expressed himself, or to some structural or stylistic aspects in the text that need to be traced. The error compels us to reconsider the text itself, and conduct a better and highly informed close reading of it, one that intends to uncover its hidden aspects.

Another conclusion that can be derived is that contrary to one's expectations, mentally disturbed or psychologically challenged individuals are portrayed with less interest in death and fear of entrapment in a world where they cannot live, and more with an acute sense of self awareness that despite its irrationality, is pretty much the product of a reasoning process that differs little from what goes on in the mind of a sane and psychologically sound individual.

Reflecting on the classification results demonstrated previously should be grounded in the fact that one cannot generalize these results (or their meanings) on other unexplored literary datasets, because approaching literary text mining tasks with predetermined concepts about the presence or otherwise of a particular theme, will be a reiteration into the world of subjective literary analyses. However, the researcher can use the tool developed for this study; i.e. the classifier to test it on unseen fictional data which deals with mental illness or psychological themes to discover whether or not these texts reference the three topics investigated in this study, entrapment, death, and self-reflection.

References

Aggarwal, Charu C., & Zhai, ChengXiang. (2012). A survey of text classification algorithms. *Mining Text Data*. Springer US, pp 163-222.

Aggarwal, C. C., & Zhai, C. (2012). An introduction to text mining. In *Mining text data* (pp. 1-10). Springer US.

Barufaldi, B., Santana, E. F., Filho, J. R., Poel, J. V., Marques, M. & Batista, L. V. (2009). Text classification by literary period using PPM-C Data Compression. *Seventh Brazilian Symposium in Information and Human Language Technology, Sao Carlos, TBD, Brazil*, pp. 125-133.

Blevins, Cameron. (April 1, 2010). Topic modeling Martha Ballard's Diary. *Historying*.

Brownlee, J. (2016). How to estimate the performance of machine learning algorithms in Weka. *Weka Machine Learning*. Retrieved from [http://machinelearningmastery.com/estimate-performance-machine-learning-algorithms-weka/]

Ciobanu, A. M., Dinu, L. P., Sulea, O. M., Dinu, A., & Niculae, V. (2013). Temporal text classification for Romanian novels set in the past. *RANLP* (pp. 136-140).

Clement, T. (2013). Text analysis, data mining, and visualizations in literary scholarship. *Literary Studies in the Digital Age: An Evolving Anthology*.

Cranenburgh, A., Huygens, I. N. G., & Koolen, C. (2015). Identifying literary texts with bigrams. *Computational Linguistics for Literature*, 58

Donnelly, T. (2012). Vogue diagnoses: Functions of madness in Twentieth-Century American literature.

Flanders, Julia H. (2009). Digital humanities and the politics of scholarly work. Diss. Brown University.

Giang, B. T. H. (2014). Trauma and psychological losses in F. Scott Fitzgerald's novels (Doctoral dissertation, Fukuoka Women's University).

Harper, S. (1997). The subject of madness: insanity, individuals and society in late-medieval English literature (Doctoral dissertation, University of Glasgow).

Hettinger, L., Jannidis, F., Reger, I., and Hotho, A. (2016). Significance testing for the classification of literary subgenres. *Digital Humanities: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 218-220.

Höhn, M. E. E. (2007). Out of her mind: female insanity from the 1890s to the 1970s on the basis of women's social history in" The yellow wallpaper", The bell jar, and Surfacing (Master's thesis).

Jockers, M. L. (2013a). Testing authorship in the personal writings of Joseph Smith using NSC classification. *Literary and linguistic computing*, 28(3), 371-381.

Jockers, M. L. (2013b). Macroanalysis: Digital methods and literary history. University of Illinois Press.

Jockers, M. L., & Mimno, D. (2013). Significant themes in 19th-century literature. Poetics, 41(6), 750-769.

Jose, L. (2010). Madness and gender in late-medieval English literature (Doctoral dissertation, Durham University).

Liu, A. (2012). The state of the digital humanities: A report and a critique. *Arts and Humanities in Higher Education*, 11(1-2), 8-41.

Love, H. (2010). Close but not deep: Literary ethics and the descriptive turn. New Literary History, 41(2), 371-391.

Moretti, Franco. (2000). Conjectures on world literature. *New Left Review* 1, 54–68.

Ramsay, S. (2003). Special section: Reconceiving text analysis toward an algorithmic criticism. *Literary and Linguistic Computing*, 18(2), 167-174.

Rieger, B. M. (Ed.). (1994). Dionysus in literature: essays on literary madness. University of Wisconsin Pres.

Svensson, Patrik. (2010). The landscape of digital humanities. *Digital Humanities Quarterly*, 4(1). Retrieved from [http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html]

Thiher, Allen. Revels in madness: Insanity in medicine and literature. Ann Arbor: U of Michigan P, 1999.

Yu, B. (2008). An evaluation of text classification methods for literary study. *Literary and Linguistic Computing*, 23(3), 327-343.