International Journal of Modern Agriculture
ISSN: 2305-7246
Volume 10 Issue 3, 2021

CLASSIFICATION TECHNIQUES OF SENTIMENTS AT DIFFERENT ANALYSIS LEVELS IN CONTEXT OF OPINION EXTRACTION: A REVIEW

Mrs.K.Balasaranya[1], Dr.P.Ezhumalai[2]

[1] Assistant Professor, Department of CSE, R.M.D. Engineering College, Kavaraipettai, India
[2] Professor and Head, Department of CSE, R.M.D. Engineering College, Kavaraipettai, India
Email: [1]balasaranya1701@gmail.com, [2]ezhumalai.es@gmail.com

Abstract

In recent days, web marketing and digital environment is pervading everywhere for fulfilling the predominant needs of people especially in Covid-19 pandemic. In general, human beings are strongly associated with sentiments and others opinion. Sentiment analysis contributes a major role in mining the opinions and views of the people. Opinion extraction facilitates the ranking process of search results in an optimal manner. Opinion categorization is done mainly using polarity strength that is obtained through sentimental analysis. This necessitates the implementation of sentiment analysis techniques at Document Level, Sentence Level and Aspect Level. Document level based sentimental analysis calculates the opinion dimensional strength by analyzing the whole document. However, opinion strength of intended domain of the search is determined by applying Sentence level based sentimental analysis. At intrinsic level, to achieve opinion strength of the features or aspects, Aspect level based sentimental analysis is devised. Sentimental techniques are expended in classifying the intentions of users at all levels that facilitates ranking process. This paper examines the classification techniques in all analysis levels. Also summarizes the research strategies enforced for classification, their strengths and captivates the scope of development.

Key words: Opinion Categorization, Opinion Mining, Polarity Strength, Ranking, Sentiment Analysis, Sentiment Classification Levels

Introduction

Human minds and actions are frequently connected by emotions. In general, the tendency of humankind is to affiliate with sentiments. Nowadays, a strong conventional association subsists between the online commercial activities and sentiments of the people. In addition, today's activities are bounded to time. People are busy in their schedule and they do not have much of time in gathering information and finding decision. Sentimental classification leads to analyze sentiments grounded on the nature and views of the user. Analyzed result sets are utilized to draw out the user's interest and opinions. Accordingly, the search results will be concise and focused. Classification is accomplished through the excerpted aspects of people's emotions, views and opinions. This paper reviews some prevailing techniques in sentiment emotional classification in context of opinion extraction. This also covers the detailed comparison between various methods of analysis, working principle, advantages and disadvantages. The remaining portion is formulated with necessary explanations. Section II covers implication of emotional analysis in opinion mining. Section III incorporates the assorted levels of psychological categorization and segregation purpose at respective levels. Section IV addresses the pre-formulated methods used in evaluation for classifying persuasions. Section V delineates the methodologies with the comparative study and summary of analysis carried out for the techniques examined for sentimental classification. Advantages and shortcomings of reviewed techniques are highlighted. Section VI concludes the study with the discussion of some future directions for research.

Importance and Background

Sentiment analysis reposes sixth sense applied science that works by incorporating various subject fields such as Natural-Language Processing, Text Mining, Artificial Intelligence, Computational Intelligence and lot more linguistics technologies. It exploits the rationale of identifying the opinions, emotions of users computationally and segregating their emotions as sentiments. Opinions may be expressed through stimulations gathered from several sources of entropy such as surveys, discussion forums, reviews, feedback, comments, recommendations, evaluations, Web based social networks, blogs. User's views about any aspect may fall in any of these categories: Positive, Negative, or Neutral. An instance for teacher's emotion tagging on online teaching pen tablet usage is indicated in Fig.1.

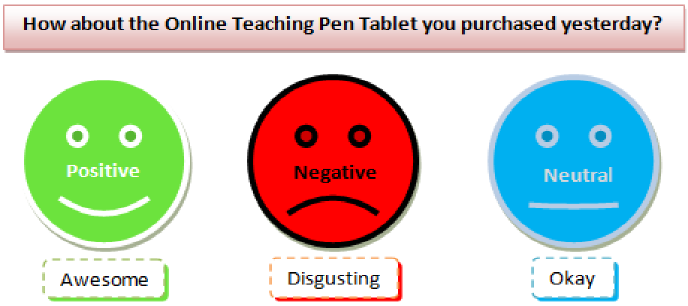

Fig . 1. Emotion Categories-Example

For commercial strategies, Sentiment evaluation is considered as contextual mining technology that helps to identify, analyze and understand the people's mentality about the vendor service, product quality, their popularity, brand reach ability which improves their service and business growth. Sentiment analysis cares for intrinsic feelings of the people. In general, People refer to customers, client, public users, or intended end users that depend on application.

Sentiment examination screens out the views, feelings or opinions, and then extracts the semantic polarity of imparted sentence. This semantic polarity aids to fix polarity strength of the assertion such as weakly positive, weakly negative, little positive, little negative, strongly positive, strongly negative, or neutral. Sentiment scrutiny judges views and emotions of the people. This aids in assessing people and their opinions. Sentimental analysis embodies people's enthusiastic opinions, expectations, poor responses; neutral feedbacks are analyzed and represented. It is useful to forecast people's future reactions and their impact towards the product or document or industry. This Predictive analysis about the future prefiguring will help the business people to equip their business in the right direction and tailor their methodologies, business rules towards the positive commercialization.

Archived views of the people, communities, or groups will also help the new users to take decisions and understand about the product or document. Information archiving strategies and retrieval techniques assist the entire process to reinforce opinions of old customers, which gives illusion to new customers concerning the product or document.

Sentimental Classification

R. Sharma et.al [1] suggested that there exist three basic levels for Sentiment mining or Classification process: 1. Document-Level Classification 2. Sentence-Level Classification 3. Aspect-Level Classification. Sometimes, Aspect level of classification replaces Feature level Classification. Fig.2 shows the hierarchy of classification levels of analysis.

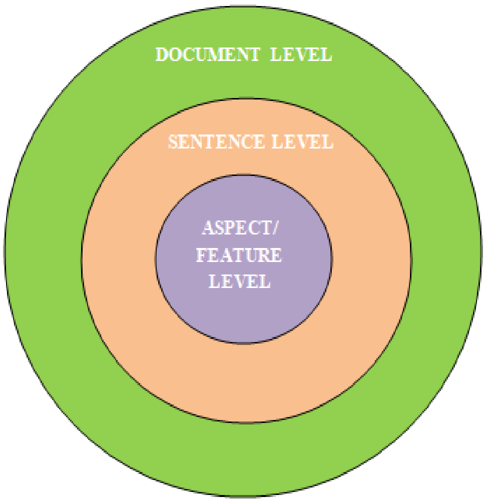

Fig. 2. Levels of Sentimental Classification

Review of Techniques

This section wields the diverse methods on thoughts and sentiment categorization used for mining of opinions at three levels: Document, Sentence, Aspect or Phrase. In all the three levels, some commonly used techniques are identified, studied in detail, their related aspects are summarized.

A. Document Level Techniques

Document based sentiment analysis focuses on examining of sentiment judgment of entire document. Document concerns to any data source in section I. This level insures the polarity for document and categorizes the document with overall polarity strength. In document level, each document addresses single entity of data source. This section gives the review for document level classification techniques.

Chen, L. S et.al [2] explored the method of Artificial Neural-Network (ANN) for determining the index orientation metrics over document semantics. The extracted orientations are mapped to regulate the polarity level with Self-organizing Maps (SOM). Lexicon sets of Sentiments over feature choice is rarely found.

Anuj Sharma et.al [3] proposed the method called Back-Propagation Artificial Neural-Network (BPANN) for analysis at document level. It combines machine learning schemes and lexicon tailored classification methods. BPANN uses the principle of mining the sentiment features from the original data source at document level. The resulted features are given for special layers in BPANN which analyses and provide a non-linear function. This complex non- linear function identifies a sentiment polarity pattern that classifies the document.

Sentiment oriented lexicons are evoked with information gain measurement from Bag-of-Words of document. The distilled lexicons are analysed in detail to represent sentiment related features which are integrated as training or test datasets. Intrinsic cognition based on semantics and sentiments are extracted as lexicons.

Jawad Khan et.al [4] proposed the technique of Ensemble-Learning schemes for document phase of sentimental classification. This addresses four major issues in document analysis. (1) Feature Extraction: Appropriate features are investigated and mined to build lexicons for sentiment classification. (2) Dimensionality Reduction: Extracted feature space are reduced by choosing the meaningful subsets which comprises effective aspects alone. (3) Ensemble Model: From the Dimensionality space, needed characteristics are elicited over filter based selection techniques. With the help of majority-voting schemes, interested aspects are selected and the document is classified accordingly. (4) Practicality: Authentication and testing is provided over standard blogs, reviews and other document sources.

B. Sentence Level Techniques

Sentiment classification for examining the sentences focuses the psychological emotions of user at each sentence. Factual opinions are uttered as opinions. Each sentence is examined for its polarity level and polarity strength is calculated and accordingly the viewpoints are classified.

Duyu Tang et.al [5] in his research, proposed the method of joint strategy and framework of sentence segmentation and classification. It employs the principle of producing useful segmentations for each sentence. At each segment level, sentence polarity is analysed and predicted concurring to segment results. Candidate-Generation framework generates segmentation tailored candidates of each sentence. Using segmentation ranking exemplary, usefulness scores of segmentation candidate are estimated. Sentiment Polarity is calculated using predictive classification model.

Wei Zhao et.al [6] implemented a deep learning approach for categorizing sentences. It exploits weak supervision signals to classify. This approach embeds the layer of deep learning canvas on the top of sentiment-distribution rating layer. Deep embedding works in two stages: (1) Embedding domain that comprises of sentimental distribution details of sentences is created. By supplying rating knowledge of sentences gathered from the source document, distribution ranges are finalized. This embedding space gives the illusion of high level indication and intrinsic features are extracted. (2) On the top of embedding space, another layer called classification layer is introduced which uses labeled sentences which helps in fine tuning the result of supervised sentence classification. Convolution features evoke the thoughts and polarization strength is assigned.

C. Aspect Level Techniques

In aspect category of sentiment representation, recognition scope of negation words are acknowledged. This enables the sentimental classification in an appropriate feature categorization and analysis At the Aspect or feature level mining process, evaluation schemes and correlated techniques are enforced for perspective level assessment grouping and classification.

At the Aspect or feature level mining process, some valuations are exercised which is called perspective level assessment grouping and classification.

Zhengjie Gao et.al [7] proposed a variation of Bidirectional-Encoder Representation Transformer (BERT) to perform NLP based aspect classification. This model ferments the target information into features. Target knowledge is distilled from the set of features that are embedded in document or sentence. Appropriate sentence level analysis is induced and sentences are classified with polarity strength. The result is protruded in a special layer called Fully-connected layer which is called BERT-FC. It also incorporates input from target dependent information. In the situation where there is multiple target information, a Max pooling method is executed as opening step to constrict the layer. Then, feasting of data to the upcoming fully-connected layer is accomplished. Relevant and effective tokenizer is used for further processing.

Aishwarya Mohan et.al [8] devised an algorithm called Sentiscore algorithm which automatically calculates the sentimental scores of aspects with principal user's opinions. Sentiscore algorithm embodies aspect-level classification of opinions founded on sentiment factor. As the opening move, Data Cleansing was carried out. Then prediction of polarity was achieved using Rule-based classifier which uses priority based algorithms. Using Sentiscore algorithm, aspect level document matrix was generated. Based on the matrix orientation, further refinement was done which gives inferences about the decisions to be made.

Summary and Findings

Some prevailing techniques in sentimental classification that is adherent to opinion centered mining were studied. Table I captures the summary of the illations of detailed study performed at all the levels of classification.

Table I: Summary and Inferences of Techniques Reviewed

Technique Used	Level of Sentimental Classification	Paper Reference Number	Targeted Domain Analyzed	Examples for Targeted Domain Data Sources	Advantages	Disadvantages
Artificial Neural Network	Document Level	[2]	Public Discussions	• Cluster Web blogs	• Polarity Strength identification is optimum • Exact results for smaller dataset	• Minimal involvement of lexicons • Feature Selection process needs more training
Back-Propagation ANN	Document Level	[3]	Consumer Reviews	• Data Corpus from Movie Reviews • Data sets from Hotel Reviews	• Reduced Dimensionality adaptive learning • Parallelism • Pattern learning • Sequence recognition • Fault tolerance	• Non-convergence • Overfitting issues • Sensitive to training sets
Ensemble Learning	Document level	[4]	Business Recommen dations	• Reviews for Commercial Products • People Opinion blogs	• Classification accuracy is high • Reduced dimensionality feature space • Relevant Sentiment-bearing features	• Idioms are not extracted as lexicons • Metaphors, irony terms are not addressed • Not dealing with Sarcasm
Joint Segmentation	Sentence level	[5]	Social Networking	• Tweets • Review Ratings	• Polarity Annotations of sentence is good • Simultaneous process of generation of segments and sentence level prediction is achieved	• Distant Supervision is weak
Deep Embedding technique	Sentence level	[6]	Contextual Recommen dations	• Review sentences • Merchant sites Review sites	• Stronger Expressive power	• Large scale training data is needed • Labeling is expensive
Bidirectional Encoder-Representation s from Transformers	Aspect level	[7]	Experiment al statistics	• Feedbacks from E-Commerce Sites	• Context based polarity strength is good • Fine-tuned classification accuracy	• Need for Pre-trained embeddings of words • Semantic information analysing in the global context is limited
Sentiscore Algorithm	Aspect level	[8]	Customer Reviews	• Restaurant Domain	• Easy to fix up opinion polarity • Aspect matrix provides good strength • Positive instances are more when k-nn classifier is used. • Easily detects Polarity of Multi-aspect sentence	• Sarcasm within reviews are not handled • Timestamp is not an attribute for classification

Some standard Evaluation Metrics of opinions are considered and analyzed with various level of sentimental classification. Impact and correlating level of metric is inquired with the help of prevailing research findings [9-17]. Accuracy levels of the metrics are examined and figured out in percentage at all the levels of classification. Table II renders the research findings of above said researchers in accuracy level perspective of evaluation metrics.

Consorting the percentage of resulted accuracy concluded by the existing explorations [9-17], the impact correlation percentage of standard evaluation metrics for each level of sentimental classification are visualized in Figure 3.

Table II. Accuracy Percentage of Metrics at different Classification Levels

Evaluation Metric	Accuracy (in terms of Percentage)		
	Document Level	Sentence Level	Aspect Level
Precision	21.63	42.09	56.79
Recall	56.89	71.47	75.02
F-Score	12.31	13.46	41.87
Polarity Strength	46.91	73.66	81.73
Adherence	22.52	67.81	42.69
Mapping Strength	45.09	71.64	92.31
Augmentation	56.72	45.71	67.36

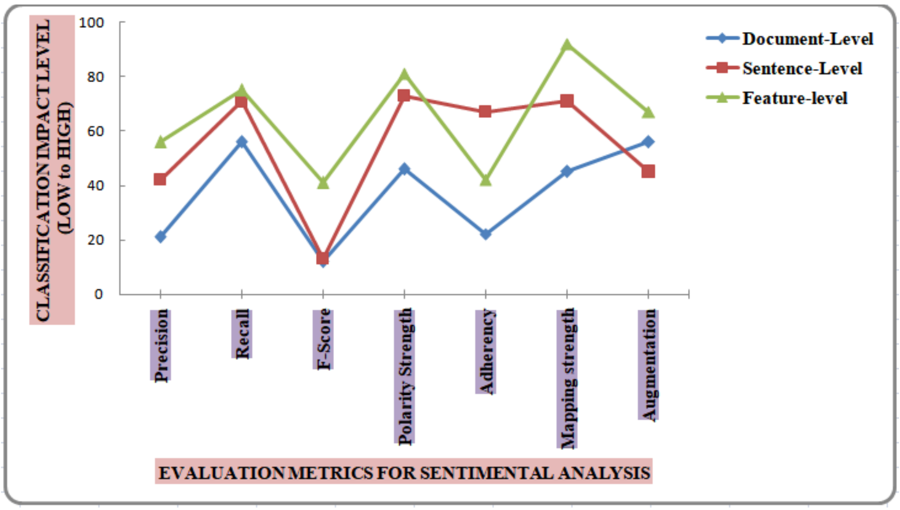

Fig 3. Accuracy Levels of Evaluation Metrics

In accordance to above findings, the impact level of correlation of basic metrics are reviewed and interpreted. Grounded on the accuracy percentage shown above, the impact level is catered as low, moderate and high as given in the Table III.

Table III. Impact Level of Evaluation Metric at different Classification Levels

Evaluation Metric	Document Level	Sentence Level	Aspect Level
Precision	Low	Moderate	Moderate
Recall	Moderate	High	High
F-Score	Low	Low	High
Polarity Strength	Moderate	High	High
Adherence	Low	Moderate	Moderate
Mapping Strength	Moderate	Moderate	High
Augmentation	Moderate	Moderate	High

Conclusion and Scope for Future Work

Standard techniques implicated for processing sentimental classification at distinct stages of analysis are reviewed. These techniques are useful to tap feelings and notions of end-user or customer. A thorough study brought to understand emotions of users and its impingement in the polarity setting. Different techniques are explored to probe the feedbacks and reviews recorded by the user for improving the concert of business application or product. Extracting Document level Polarity and obtaining the sentence polarity plays substantial role in sentimental categorization of the feature taken in broader view. Narrowing down our context where opinions enrich digital marketing and ranking, polarity of sentiment aspects are necessary. Elaborative reviews were put through in all levels in view of opinion mining. Summary of this work is tabulated with methodology incorporated, targeted data source used, merits and disadvantages. This summarization leads the researcher to augment their research in this sentimental classification that results in potential algorithms and techniques.

With relevant techniques reviewed and detailed study carried out, it is inferred that all methodical levels must be precisely analyzed to obtain the optimum categorization. In future, Bio-Inspired algorithms can be employed over feature level aspects. This may result in more adherent opinion extraction. Today's digital world relies on favorable search results and ranking. Due to Covid-19, everything becomes digital, which makes dynamic ranking as an immense confront. As a part of future scope, dynamic ranking algorithms can be designed for incurring the desired search results of opinion extraction process.

References

1. R. Sharma, S. Nigam, R. Jain, "Polarity detection at sentence level", International Journal of Computer Applications, vol. 86, pp.11-21, 2014.
2. Chen L.S, Liu C.H, Chiu H.J, "A Neural Network based approach for sentiment classification in the blogosphere", Journal of Informatics, pp.313-322, 2011.
3. Anuj Sharma, S.Dey, "A document-level sentiment analysis approach using artificial neural network and sentiment lexicons", ACM SIGAPP Applications Comput. Rev., vol. 12, no. 4, pp. 67–75, 2012.
4. Jawad Khan, Aftab Alam, Jamil Hussain, Young-Koo, "EnSWF: effective features extraction and selection in conjunction with ensemble learning methods for document sentiment classification", Appl. Intell., vol. 49, no. 8, pp. 3123–3145, 2019.
5. Duyu Tang, Bing Qin, Furu Wei, Li Dong, Ting Liu, Ming Zhou, "A Joint Segmentation and Classification Framework for Sentence Level Sentiment Classification," IEEE Transactions on Audio, Speech Language and Processing, vol. 23, No. 11, pp.1750–1761, 2015.
6. Wei Zhao, Ziyu Guan, Long Chen, Xiaofei He, Deng Cai, Beidou Wang, Quan Wang, "Weakly-supervised deep embedding for product review sentiment analysis", IEEE Transactions on Knowledge and Data Engineering, vol.30, No.1, pp.185–197, 2018.
7. Zhengjie Gao, Ao Feng, Xinyu Song, Xi Wu, "Target-dependent sentiment classification with BERT", IEEE Access, vol.7, pp.154290–154299, 2019.
8. Aishwarya Mohan, Manisha.R, Vijayaa.B, Naren.J, "An Approach to Perform Aspect level Sentiment Analysis on Customer Reviews using Sentiscore Algorithm and Priority Based Classification Expert Systems in Affective Computing View project Gateman View project An Approach to Perform Aspect level Sentiment", Int. J. Comput. Sci. Inf. Technol., vol. 5(3), no. August, pp. 4145–4148, 2014.
9. Claster W.B, Dinh Quoc H, Shanmuganathan S, "Unsupervised Artificial Neural Nets for Modeling Movie Sentiment", Proceedings of Second International Conference on Computational Intelligence, Communication Systems and Networks (CICSyN'10), Beppu, Japan, pp349-354, 2010.
10. D.Krishna Madhuri ,"A machine learning based framework for sentiment classification: Indian railways case study ", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 4, pp. 441–445, 2019.
11. P. Tyagi, R. C. Tripathi "A Review towards the Sentiment Analysis Techniques for the Analysis of Twitter Data ", SSRN Electron Journal, doi: 10.2139/ssrn.3349569 ,2019.
12. A. Alsaeedi, M. Z. Khan ,"A study on sentiment analysis techniques of Twitter data ", International Journal of Advanced Computer Science and Applications, vol. 10, no. 2, doi: 10.14569, pp. 361–374 , 2019.
13. M. Z. Asghar, A. Khan, S. Ahmad, M. Qasim, I. A. Khan , " Lexicon-enhanced sentiment analysis framework using rule-based classification scheme ", PLoS One, Vol.No: 12, doi: 10.1371/journal.pone.0171649, pp. 1–22 , 2017.

14. J. Singh, G. Singh, R. Singh, “Optimization of sentiment analysis using machine learning classifiers”, Human-centric Computer and Information Science, vol. 7, no. 1, doi: 10.1186/s13673-017-0116-3, 2017.
15. K. Sarvabhotla, P. Pingali,V. Varma ,“Sentiment classification: A lexical similarity based approach for extracting subjectivity in documents ”, Information Retrieval, Boston., vol. 14, no. 3, doi: 10.1007/s10791-010-9161-5, pp. 337–353 , 2011.
16. A. Jain, B. Pal Nandi, C. Gupta , D. K. Tayal, “Senti-NSetPSO: large-sized document-level sentiment analysis using Neutrosophic Set and particle swarm optimization ”, Soft Computing, vol. 24, no. 1, doi: 10.1007/s00500-019-04209-7, pp. 3–15, 2020.
17. B. S. Rintyarna, R. Sarno, C. Fatichah, “ Evaluating the performance of sentence level features and domain sensitive features of productreviews on supervised sentiment analysis tasks ”, Journal of Big Data, vol. 6, no. 1, doi: 10.1186/s40537-019-0246-8 , 2019.
18. J. Zhou, J. X. Huang, Q. Chen, Q. V. Hu, T. Wang , L. He,“ Deep learning for aspect-level sentiment classification: Survey, vision, and challenges ”, IEEE Access, vol. 7, doi: 10.1109/ACCESS.2019.2920075, pp. 78454–78483, 2019.
19. E. Cambria, D. Das, S. Bandyopadhyay , A. Feraco, “Affective computing and sentiment analysis ”, In A practical guide to sentiment analysis, ed: Springer, pp. 1-10 ,2017.
20. H. Keshavarz, M. S. Abadeh,“ALGA: Adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs ”, Knowledge-Based Systems, vol. 122, pp. 1-16, 2017.
21. Y. Song, J. Wang, T. Jiang, Z. Liu ,Y. Rao , “Attentional encoder network for targeted sentiment classification ”, arXiv preprint arXiv:1902.09314 , 2019.
22. A. Lu, W. Wang, M. Bansal, K. Gimpel , K. Livescu , “ Deep multilingual correlation for improved word embeddings ”, Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 250-256 , 2015.
23. D. Osimo , F. Mureddu, “Research challenge on opinion mining and sentiment analysis” ,*Univ. Paris-Sud, Lab. LIMSI-CNRS, B{â}timent*, vol. 508, 2012.
24. D. Mohey El-Din, “Analyzing Scientific Papers Based on Sentiment Analysis (First Draft) Sentiment Analysis of Online Papers (SAOOP) View project Egypt Blockchain View project” , no. April, doi: 10.13140/RG.2.1.2222.6328 , 2016.
25. X. Pu, G. Wu, C. Yuan, “Exploring overall opinions for document level sentiment classification with structural SVM,” *Multimed. Syst.*, vol. 25, no. 1, pp. 21–33, doi: 10.1007/s00530-017-0550-0 , 2019.
26. Luke Kien-Weng Tan, Jin-Cheon Na, Yin-Leng Theng, Kuiyu Chang, "Phrase-Level Sentiment Polarity Classification Using Rule-Based Typed Dependencies and Additional Complex Phrases Consideration", Journal Of Computer Science and Technology, 27(3): 650–666 May 2012.