University of East London

Intelligent Voice

STRENUUS®

Fraud detection in telephone conversations for financial services using linguistic features

Nikesh Bajaj[1], Tracy Goodluck Constance[1], Marvin Rajwadi[1], Julie Wall[1], Mansour Moniri[1], Cornelius Glackin[2], Nigel Cannings[2], Chris Woodruff[3], James Laird[3]

[1] Intelligent Systems Research Group, University of East London, [2] Intelligent Voice Ltd. [3] Strenuus Ltd.

Introduction

- Detecting element of deception is challenging and a major concern for safe and trustworthy communication.
- The majority of work in fraud detection utilise banking transactions and activities.
- A telephone conversation is usually the first point of contact for a customer, which can contain potential fraudulent cues at a very early stage to prevent fraud.
- We propose an approach to use linguistic features based on the semantic and syntactic structure of the transcribed conversation.
- We validate our approach with real-world financial services dataset.

Proposed approach: Linguistic Features

In literature, many linguistic markers are defined to indicate the emotional and cognitive state of a speaker. Together with linguistic and expert interrogators, we complied and derived a list of markers and their trigger terms.

Linguistic Markers

- **Causation:** Providing a certain level of concreteness to an explanation. Ex: Because, Effect, Hence
- **Negation:** Avoiding to provide a direct response [1] No, Not, Can't, Didn't
- **Hedging:** Describes words which meaning implicitly involves fuzziness [2] May be, I guess, Sort of
- **Qualified assertions:** Unveils questionable actions [2] Needed, Attempted
- **Temporal Lacunae:** Unexplained lapses of time [2] Later that day, Afterwards
- **Overzealous expression:** Expresses some level of uncertainty [2] I swear to God, Honestly
- **Memory loss:** Feigning memory loss [2] I forget, Can't remember
- **Third person plural pronouns:** Possessive determiners to refer to things or people other than the speaker [7] They, Them, Theirs
- **Pronouns:** Possessive determiners to refer to the speaker by overemphasising their physical presence [7, 5] I, Me, Mine
- **Negative emotion:** Negative expressions in word choice [8, 7, 3] Afraid, Sad, Hate, Abandon, Hurt
- **Negative sentiment:** Negative emotional effect [3] Abominable, Anger, Anxious, Bad
- **Positive emotion:** Positive expressions in word choice [7, 3] Happy, Brave, Love, Nice, Sweet
- **Positive sentiment:** Positive emotional effect [3] Admire, Amazing, Assure, Charm
- **Disfluencies:** Interruption in the flow of speech [7] Uh, Um, You know, Er, Ah
- **Self reference words:** Deceivers tend to use fewer self-referencing expressions [2] My, As I said.
- **Nominalised verbs:** Nouns derived from verbs. Nominalisations tend to hide the real action. [4] Education, Arrangement

Sentiment

- Linguistic markers analyse only the syntactic information, prone to miss the overall sentiment.
- The sentiment reflects the polarity of the speaker's feelings in a dimension from negative to positive.
- A Deep Neural Network, trained on IMDB dataset [6] is used to extract the sentiment values of customer's responses.

Experiment

Dataset

- Real-world data collected from financial services institutions
- Transcriptions of 56 telephone conversations, 32 Fraud, 24 Non-Fraud

Feature Extraction and Modeling

- From each conversation, the frequency of each linguistic marker is computed (16 features).
- The sentiment of each response of the customer is estimated using pre-trained model. A total 11 statistical measures from collected sentiment values for each conversation are computed. Statistical measures are namely: Mean, SD, Min, Max, Median, IQR, Kurtosis, Skewness, Positive energy (pE), Negative energy (nE), and total number of responses (tR).
- Four classifiers namely: Naive Bayes, Decision Tree (DTree), k-Nearest Neighbors (kNN), and Support Vector Machines (SVM) are trained with individual features first then with combined features.
- Th parameters for classifiers are as follow; Decision Tree with depth 3, kNN with 3 neighbours, and SVM with linear kernel.
- Each model is trained and tested with K-Fold cross-validation, with K=10. The mean and SD of the training and testing accuracies are tabulated in Table 1 and plotted in Figure 2.

Results

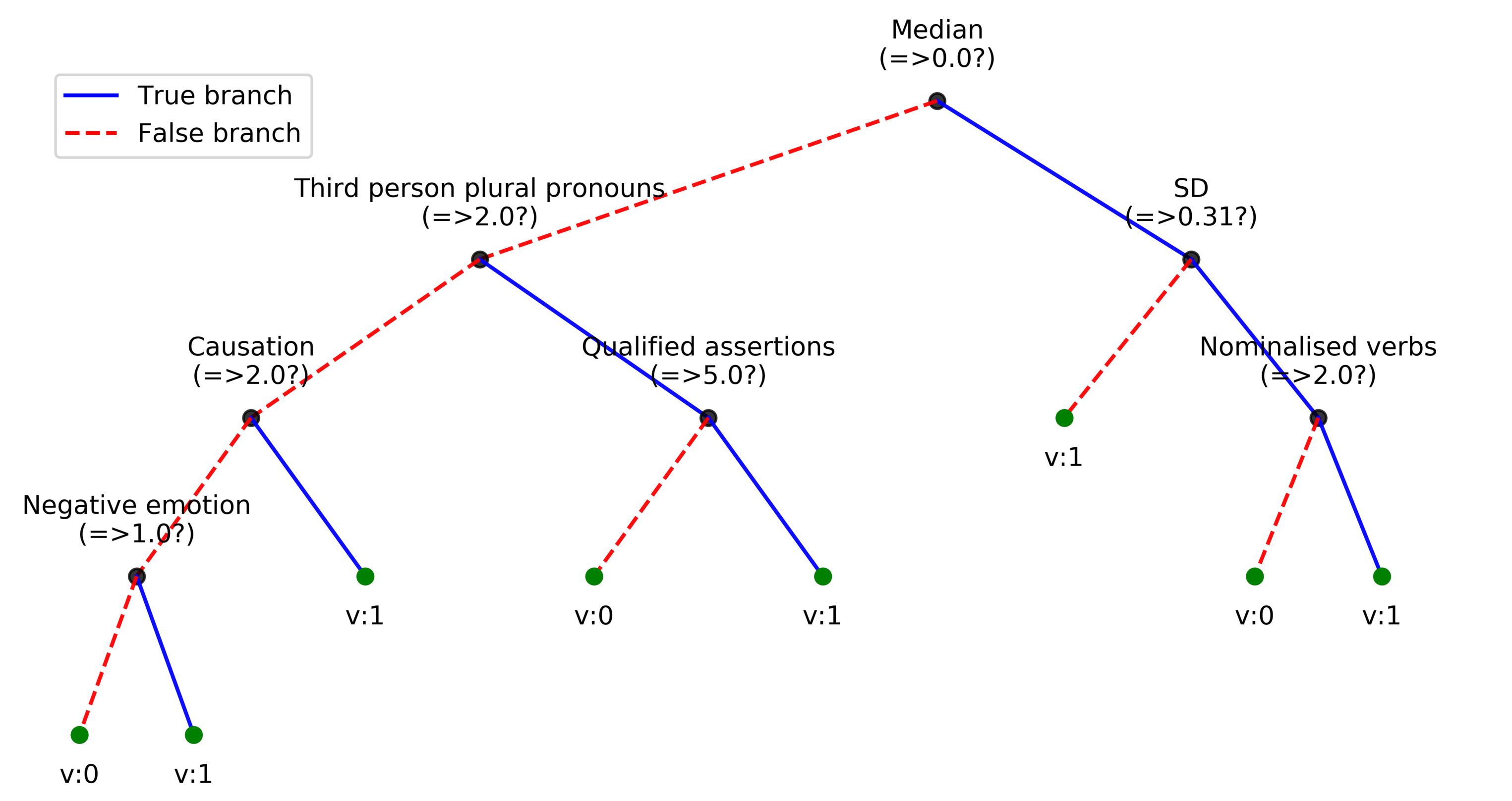

Figure 1: A Decision Tree for fraud detection. Leaf node v:0 - Non-Fraud, v:1 - Fraud

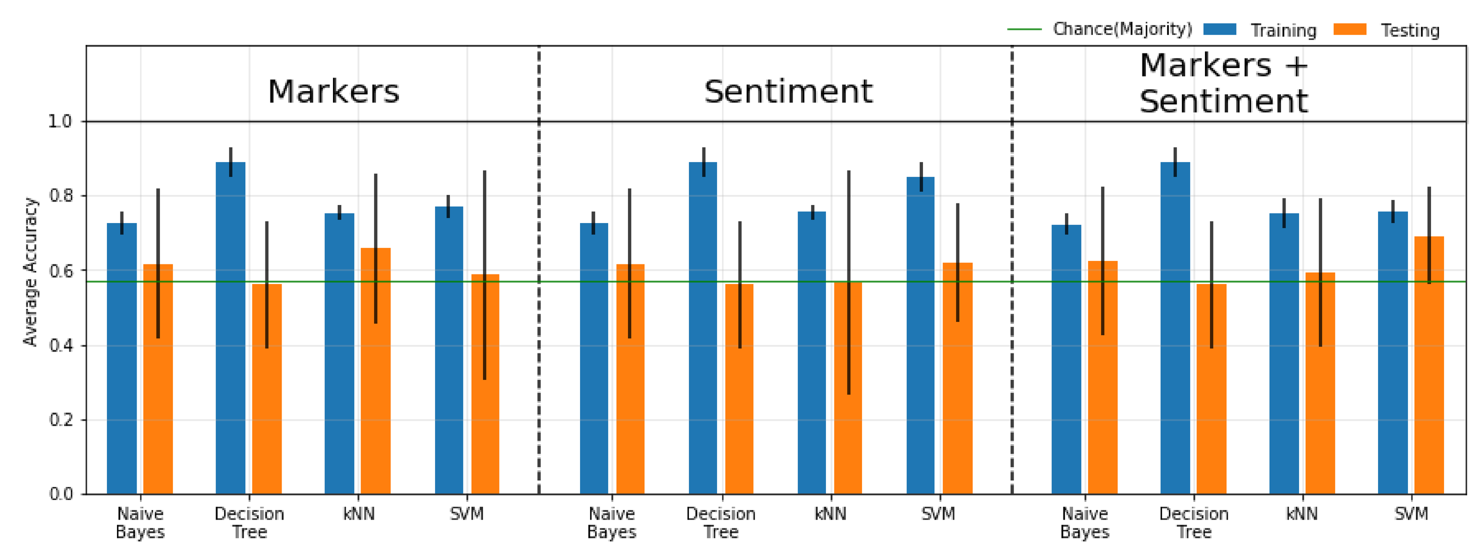

Figure 2: Average performance of K-Fold(K=10) for different models

Table 1: Results of modeling with K-Fold (K=10)

Features	Accuracy	Model			
		Naive Bayes	DTree (d=3)	kNN(k=3)	SVM(Linear)
Markers	Training	0.73 ± 0.03	0.89 ± 0.04	0.75 ± 0.02	0.77 ± 0.03
	Testing	0.62 ± 0.20	0.56 ± 0.17	0.66 ± 0.20	0.59 ± 0.28
Sentiment	Training	0.72 ± 0.03	0.89 ± 0.04	0.75 ± 0.02	0.85 ± 0.04
	Testing	0.62 ± 0.20	0.56 ± 0.17	0.57 ± 0.30	0.62 ± 0.16
Markers + Sentiment	Training	0.72 ± 0.03	0.89 ± 0.04	0.75 ± 0.04	0.76 ± 0.03
	Testing	0.62 ± 0.20	0.56 ± 0.17	0.59 ± 0.20	0.69 ± 0.13

- The highest testing accuracy achieved with solely the linguistic markers is **65.5%** using kNN, whereas using sentiment features it is **62%** using SVM. However, combined features improve the accuracy to **69%** using SVM.
- Decision Tree shown in Figure 1, suggests the median of sentiments values is more important than others.

Conclusion & Future work

- Results suggest the proposed approach has the potential to exploit the linguistic features to detect the fraud and deception in a transcribed conversation. Simple models can be employed for real-time process and utilise to explain the decision process.
- Future work plans to extend the presented work for different scenarios including legal and insurance services, to again employ real-world data.

References

[1] S. H. Adams. *Communication under stress: indicators of veracity and deception in written narratives.* PhD thesis, Virginia Tech, 2002.

[2] J. Bachenko, E. Fitzpatrick, and M. Schonwetter. Verification and implementation of language-based deception indicators in civil and criminal narratives. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 41–48. Association for Computational Linguistics, 2008.

[3] D. Jurafsky. *Lexicons for Sentiment, Affect, and Connotation.* Pearson Education India, 2000.

[4] M. Lapata. The disambiguation of nominalizations. *Computational Linguistics*, 28(3):357–388, 2002.

[5] J. Li, M. Ott, C. Cardie, and E. Hovy. Towards a general rule for identifying deceptive opinion spam. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1566–1576, 2014.

[6] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts. Learning word vectors for sentiment analysis. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1*, pages 142–150. Association for Computational Linguistics, 2011.

[7] J. W. Pennebaker, M. E. Francis, and R. J. Booth. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001, 2001.

[8] D. B. Skillicorn and A. Little. Patterns of word use for deception in testimony. In *Security Informatics*, pages 25–39. Springer, 2010.

Download full paper here

Funded by

Innovate UK
Technology Strategy Board

Contacts: {n.bajaj, j.wall}@uel.ac.uk