Publications

Journal papers

1. Lutu, P. E. N. & Engelbrecht A.P. (2013) Positive-Versus-Negative classification for model aggregation in predictive data mining, *Institute for Operations Research and the Management Sciences Journal on Computing (INFORMS JOC),* 25 (4) 792-807.

2. Lutu, P. E. N. & Engelbrecht A.P. (2013) Base model combination algorithm for resolving tied predictions for k-nearest neighbor aggregate models, *Institute for Operations Research and the Management Sciences Journal on Computing (INFORMS JOC),* 25 (3), 517-526.

3. Lutu, P. E. N. & Engelbrecht, A. P. (2012) Using OVA modeling to improve classification performance for large datasets. *Expert Systems with Applications*, 39 (4) 4358-4376.

4. Motjolopane, R. & Lutu, P.E.N. (2012) Improving data quality in the banking supervisory data of Southern Africa central banks. *Electronic Journal of Information Systems in Developing Countries (EJIDC),* 50. (also appears in the *Proceedings of the International Development Informatics Association Conference, IDIA 2010,* Cape Town, November 2010)

5. Lutu, P.E.N. & Engelbrecht, A.P. (2010) A decision-rule based method for feature selection in predictive data mining. *Expert systems with applications*, 37, 602-609.

6. Lutu, P.E.N & Engelbrecht, A.P., (2006) A comparative study of sample selection methods for classification. *SACJ: Joint Special Issue ARIMA & SACJ - Advances in End-User Data Mining Techniques*, 36, 69-85

National conference papers

1. Lutu, P.E.N. (2015) Big Data and NoSQL Databases: New Opportunities for Database Systems Curricula. *Proceedings of the South African Computer Lecturers Association Conference, SACLA 2015, July 3-4 2015, Johannesburg, South Africa. ACM International Conference proceedings.*

2. Lutu, P.E.N. (2011) Empirical comparison of four classifier fusion strategies for positive-versus-negative ensembles. *Proceedings of the South African Institute of Computer Science and Information Technology Conference, SAICSIT 2011, October 3-5 2011, Cape Town, South Africa. ACM International Conference proceedings.*

3. Lutu, P.E.N. (2007) Planning for the Assessment of Group Work for an Honours Level Course in Informatics,

South African Computer Lecturers Association Conference, SACLA 2007, Victoria Falls, Zimbabwe, July 2007.

4. Lutu,P.E.N. (2002) An integrated approach for scaling up classification and prediction algorithms for data mining. proceedings of the *South African Institute of Computer Science and Information Technology Conference, SAICSIT 2002. ACM International Conference Proceedings.*

International conference papers

1. Lutu, P.E.N. (2018) Improving Classification Performance for Minority Classes through the use of Positive-Versus-Negative Ensemble Models. *Proceedings of the 2018 IEEE International Conference on Intelligent and Innovative Computing Applications (ICONIC 2018) conference*, Mauritius, 6-7 December, 2018.

2. Lutu, P.E.N. (2018) The use of histogram analysis to support fast selection of predictive features for data stream mining. *Proceedings of the IEEE icABCD 2018 conference,* Durban, 6-7 August, 2018.

3. Nhlabano, V.V. & Lutu, P.E.N. (2018) Impact of Text Pre-processing on the Performance of Sentiment Analysis Models for Social Media Data*. Proceedings of the IEEE icABCD 2018 conference,* Durban, 6 - 7 August, 2018.

4. Lutu, P.E.N. (2017) Automated selection of training data and base models for data stream mining using Naive Bayes ensemble classification, *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering* 2017, 5-7 July, 2017, London, U.K., pp. 618-623

5. Lutu, P.E.N. (2016) Data analytics to support social media marketing: challenges and opportunities*, Proceedings of the 2016 International Conference on Information Resources Management* (*CONF-IRM 2016),* Cape Town, South Africa, 18-20 May, 2016, ISBN 978-0-473-35594-4

6. Lutu, P.E.N. (2015) Naïve Bayes classification ensembles to support modeling decisions in data stream mining, *Proceedings of IEEE SSCI 2015*, Cape Town, South Africa, 8-10 December, 2015.

7. Lutu, P.E.N. (2014) Using One-Versus-All classification ensembles to support modeling decisions in data stream mining. *Proceedings of the 3rd Workshop on New Frontiers for Mining Complex Patterns (NFMCP 2014),* collocated with the ECML-PKDD 2014 conference, Nancy, France, 15-19 September, 2014, pp. 86-97. Available at: http://www.di.uniba.it/~loglisci/NFmcp2014//program.html

8. Lutu, P.E.N. (2013) Fast feature selection for Naïve Bayes classification in data stream mining, in S.I. Ao, L. Gelman, D.W.L. Hukins, A. Hunter and A.M. Korsunsky (eds.) *Proceedings of the World Congress on Engineering (WCE 2013)*, London, U.K. July 2013, vol. III, 1549-1554.

9. Lutu, P.E.N. (2012) Generation of an empirical theory for positive-versus-negative ensemble classification. *Proceedings of the Silver 2012 Workshop, collocated with the ECML-PKDD 2012 conference,* Bristol, United Kingdom, 24-28 September, 2012.

10. Lutu, P.E.N. (2011) Using confusion matrices and confusion graphs to design ensemble classification models from large datasets, in A. Cuzzocrea and U. Dayal (eds.) *Proceedings of the 13th International Conference on Data warehousing and Knowledge Discovery, DaWak 2011,* Toulouse, France, August / September 2011, *Lecture Notes in Computer Science, LNCS 6862*, 301-315.

11. Motjolopane, R. & Lutu, P.E.N. (2012) Improving data quality in the banking supervisory data of Southern Africa central banks. *Proceedings of the International Development Informatics Association Conference, IDIA 2010,* Cape Town, November 2010.

12. Lutu, P.E.N. & Meyer, B. (2008) An exploratory study of the adoption and usage of business intelligence in public sector organisations in South Africa. *Proceedings of the IFIP WG 9.4 Workshop*, Pretoria, September 2008.

13. Adigun, E.A & Lutu,P.E.N. (2006) Mobile and wireless technology, making a world of difference. *Proceedings of the Conference on Community Informatics for Developing Countries: Understanding and Organising for a Participatory Future Information Society* (CIDC2006) 31 August -2 september 2006, Cape Town, South Africa.

14. Lutu,P.E.N. (2002) An integrated approach for scaling up classification and prediction algorithms for data mining. proceedings of the *South African Institute of Computer Science and Information Technology Conference, SAICSIT 2002. ACM International Conference Proceedings.*

15. Lutu, P.E.N. (2001) Optimal multi-splitting of numeric ranges for decision tree induction. *International Conference on Operations Research (ICORD 2001),* May 2001.

Papers presented but not published

1. Lutu, P.E.N. (2018) Graph mining methods for social networks data. Presentation at the *47th Operations Research Society of South Africa - ORSSA 2018 National Conference,* Pretoria, 16-19 September, 2018.

2. Lutu, P.E.N. (2015) Analysis of social media data: challenges and opportunities. Presentation at the *44th Operations Research Society of South Africa - ORSSA 2015 National Conference,* Johannesburg, 14-16 September, 2015.

3. Lutu, P.E.N. (2012) Replication of base models to improve the performance of positive-versus-negative classification ensembles for predictive data mining. Presentation at the *41st Operations Research Society of South Africa - ORSSA 2012 National Conference,* Johannesburg, 17-19 September, 2012.

4. Lutu, P.E.N. (2003) An Empirical Study of Classifier Accuracy in Data Mining, Operations research Society of South Africa Annual Conference (ORSSA2003), September 2003.

Book chapters

1. Lutu, P.E.N. (2006) Database sampling for data mining, in J. Wang, (editor) *Encyclopedia for Data Warehousing and Data Mining*, Information Science Reference, Hershey, New York.

2. Lutu, P.E.N. (2014), Progressive usage of business and spatial intelligence for decision support in the delivery of educational services in developing countries. In K-M Osei-Bryson and G. Mansingh (eds) *Knowledge Management for Development: Domains, Strategies and technologies for Developing Countries*. Springer 2014.

3. Lutu, P.E.N. (2015) Web 2.0 computing and social media as solution enablers for economic development in Africa. In A. Gamatié (ed.), *Computing in Research and Development in Africa: Benefits, Trends, Challenges and Solutions*, Springer International Publishing Switzerland 2014, DOI 10.1007/978-3-319-08239-4-6

4. Lutu, P.E.N. (2015) The Importance of Data Quality Assurance to the Data Analysis Activities of the Data Mining Process. In K-M Osei-Bryson and C. Barclay (eds) *Knowledge Discovery and Data Mining (KDDM) for Economic Development: Applications, Strategies and Technologies*.