INTELLIGENT CLASSIFICATION OF AMMONIA CONCENTRATION BASED ON ODOR PROFILE

MUHAMAD FARUQI BIN ZAHARI

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Master of Engineering in Instrumentation

Faculty of Electrical and Electronics Engineering
UNIVERSITI MALAYSIA PAHANG

SEPTEMBER 2016

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

INTELLIGENT CLASSIFICATION OF AMMONIA CONCENTRATION BASED ON ODOR PROFILE

MUHAMAD FARUQI BIN ZAHARI

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Master of Engineering in Instrumentation

Faculty of Electrical and Electronics Engineering
UNIVERSITI MALAYSIA PAHANG

SEPTEMBER 2016

ABSTRACT

This thesis presents the intelligent classification of ammonia concentration based on the standard of oil and gas industries wastewater discharge. The intelligent classification using signal processing is a well-known technique in many applications and as well in the oil and gas industry. The intelligent classification technique for ammonia concentration classification is a demanding technique especially in the environmental sector. Ammonia solution properties and ammonia solution preparations were studied in this thesis which commonly used in industry. The objectives of this thesis are to develop an intelligence classification of ammonia concentration based on the oil and gas industry wastewater discharge schedule and to analyze performance of the intelligent classification of ammonia concentration based on the oil and gas industry wastewater discharge schedule. In this thesis the ammonia odor profile has been pre- identified by chemist using four sensor array. The ammonia concentration was validated using a commercialized gas sensor and spectrophotometer to cross-validated e-nose instrument. The odor profile from two different samples; high (20 ppm and 25 ppm) and low (5 ppm, 10 ppm and 15ppm) concentration that have been normalized and visualized in a 2D plot to extract the unique patterns. The variance of the low and high concentration of ammonia odor profile has been identified as different group samples. This group samples have been analyzed statistically using Boxplot, calibration curve and proximity matrix. The thesis describes the statistical techniques to visualize the pattern and using mean features to classify between the low and high concentration. Two intelligent classification techniques have been used which are Artificial Neural Network (ANN) using the back-propagation approaches and then, the result of ANN model was cross-validated using CBR. Both ANN model and CBR classifier have been measured using several performance measures. From the results, it is observed that ANN model and CBR classifier are capable of classifying 100% of ammonia concentration odor profile from the water. The results can also significantly reduce the cost and time, and improve product reliability and customer confidence.

ABSTRAK

Tesis ini membentangkan klasifikasi pintar ammonia dari sisa air buangan berdasarkan standard industri minyak dan gas. Pengelasan pintar menggunakan pemprosesan isyarat adalah satu teknik yang terkenal dalam pelbagai aplikasi dan juga dalam industri minyak dan gas. Teknik pengelasan pintar untuk pengelasan ammonia adalah teknik yang ditunggu-tunggu terutama dalam sektor alam sekitar. Ciri-ciri Ammonia solution dan persediaan ammonia solution telah dikaji dalam tesis ini adalah yang biasa digunakan dalam industri. Objektif projek ini adalah untuk membangunkan klasifikasi pintar ammonia daripada sisa air buangan berdasarkan pada industri minyak dan gas mengikut pelepasan sisa air buangan yang telah ditetapkan oleh Jabatan Alam Sekitar (JAS) dan untuk menganalisis prestasi klasifikasi pintar ammonia daripada sisa air buangan berasaskan minyak dan gas industri. Dalam tesis ini profil bau ammonia telah pra dikenalpasti oleh ahli kimia menggunakan empat sensor pelbagai. Kepekatan ammonia telah disahkan menggunakan sensor gas yang telah dikomersialkan untuk silang pengesahan menggunakan E-nose. Profil bau dari dua sampel berbeza; Kepekatan yang tinggi dan rendah telah dibiasakan dan digambarkan dalam plot 2D untuk mendapatkan corak yang unik. Kepelbagaian kepekatan rendah dan tinggi bagi profil bau ammonia telah dikenal pasti sebagai kumpulan sampel yang berbeza. Ini kumpulan sampel telah dianalisis menggunakan Boxplot, Calibration Curve dan Proximity matrix sebagai teknik statistik. Tesis menerangkan teknik statistik untuk menggambarkan corak dan menggunakan ciri-ciri min untuk mengklasifikasikan antara kepekatan rendah dan tinggi. Dua teknik klasifikasi pintar telah digunakan iaitu Rangkaian Neural Buatan (ANN) menggunakan pendekatan back-propagation. Kemudian, hasil daripada model ANN adalah untuk silang pengesahan menggunakan 'Case- Based Reasoning' (CBR). Kedua-dua pengelasan pintar bagi model ANN dan pengelas CBR telah diukur menggunakan beberapa prestasi. Daripada keputusan, didapati bahawa model ANN dan CBR pengelas mampu untuk mengklasifikasikan 100% daripada profil bau kepekatan ammonia. Keputusan menunjukkan bahawa E-nose mampu mengurangkan kos dan masa ke pasaran, dan meningkatkan kebolehpercayaan produk dan keyakinan pelanggan.

REFERENCES

A. Aamodt, & Plaza, E. (1994)..Case-Based Reasoning. *Artificial Intelligence Communications, 7,* 39–59.

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications, 7,* 39–59. doi:10.1.1.56.4481

Abdallah, S. a, Al-Shatti, L. a, Alhajraf, A. F., Al-Hammad, N., & Al-Awadi, B. (2013). The detection of foodborne bacteria on beef: the application of the electronic nose. *SpringerPlus, 2,* 687. doi:10.1186/2193-1801-2-687

Adams, R., & Wales, N. (2008). NESUG 2008 Box Plots in SAS ® : UNIVARIATE , BOXPLOT , or GPLOT ?, 1–11.

Agency), (Environmental Protection. Methods for Chemical Analysis of Water and Wastes. , Pub. L. No. EPA/600/4-79/020 (1983). OHIO: United States Environmental Protection Agency.

Aguiar, A. P. D. De, Camara, G., Monteiro, A. M. V., & Souza, R. C. M. De. (n.d.). Modelling Spatial Relations by Generalized Proximity Matrices.

Akobeng, A. K. (2007). Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatrica (Oslo, Norway : 1992), 96*(3), 338–41. doi:10.1111/j.1651-2227.2006.00180.x

Aksoy, S., & Haralick, R. M. (n.d.). Feature Normalization and Likelihood-based Similarity Measures for Image Retrieval, (October 2000).

Ameer, Q., & Adeloju, S. B. (2005). Polypyrrole-based electronic noses for environmental and industrial analysis. *Sensors and Actuators B: Chemical, 106*(2), 541–552. doi:10.1016/j.snb.2004.07.033

American Society for Technical and Materials. (1972). ASTM E258: Standard Test Method for Total Nitrogen Inorganic Material by Modified Kjeldahl Method.

Astrand, M. (2008). *Normalization and Differential Gene Expression Analysis of Microarray Data.* Chalmers University of Technology and Goteborg University.

Awomeso, J. A., Taiwo, A. M., Gbadebo, A. M., & Arimoro, A. O. (2010). Waste disposal and pollution management in urban areas: A workable remedy for the environment in developing countries. *American Journal of Environmental Sciences, 6,* 26–32. doi:10.3844/ajessp.2010.26.32

Baby, R. E., Cabezas, M., Walsoe, E. N., & Reca, D. (2000). Electronic nose : a useful tool for monitoring environmental contamination, 214–218.

Bacca, J., Baldiris, S., Fabregat, R., & Avila, C. (2013). A Case-based Reasoning Approach to Validate Grammatical Gender and Number Agreement in Spanish language. *International Journal of Interactive Multimedia and Artificial Intelligence, 2*(1), 73. doi:10.9781/ijimai.2013.2110

Banerjee, S., Deka, A., Sarmah, G. D., & Bhardwaj, N. (2014). Artificial Neural Network Modeling of the Effect of Cutting Conditions on Cutting Force Components during Orthogonal Turning. *International Journal of Current Engineering and Technology,* (2), 127–130.

Barwick, V. (n.d.). Preparation of Calibration Curves : A Guide to Best Practice.

Bassey, B. E., Benka-Coker, M. O., & Aluyi, H. S. A. (2006). Characterization and management of solid medical wastes in the Federal Capital Territory, Abuja Nigeria. *African Health Sciences, 6,* 59–63. doi:10.5555/afhs.2006.6.1.58

Begum, S., Barua, S., Filla, R., & Ahmed, M. U. (2014). Classification of physiological signals for wheel loader operators using Multi-scale Entropy analysis and case-based reasoning. *Expert Systems with Applications, 41*(2), 295–305. doi:10.1016/j.eswa.2013.05.068

Bergbauer, J., Nieuwenhuis, C., Souiai, M., & Cremers, D. (n.d.). Proximity Priors for Variational Semantic Segmentation and Recognition.

Bland, J. M., & Altman, D. G. (1996). Measurement Error Proportional to the Mean. *BMJ, 313,* 106.

Bo, L., Wang, L., & Jiao, L. (2006). Feature scaling for kernel fisher discriminant analysis using leave-one-out cross validation. *Neural Computation, 18*(4), 961–78. doi:10.1162/089976606775774642

Boden, M. A. (2014). Aaron Sloman: A Bright Tile in AI's Mosaic. In J. L. Wyatt, D. D. Petters, & D. C. Hogg (Eds.), (Vol. 22, pp. 9–30). Cham: Springer International Publishing. doi:10.1007/978-3-319-06614-1

Bohm, G. (n.d.). *Introduction to Statistics and Data Analysis for Physicists.*

Broadmeadow, S., & Nisbet, T. R. (2004). The effects of riparian forest management on the freshwater environment: a literature review of best management practice. *Hydrology and Earth System Sciences.* doi:10.5194/hess-8-286-2004

Brook, T. E., Taib, M. N., & Narayanaswamy, R. (1997). Extending the range of a fibre-optic relative-humidity sensor, *39,* 272–276.

Burke, R., & Kass, A. (1996). Retrieving Stories for Case-Based Teaching. In D. B. Leake (Ed.), *Case-Based Reasoning: Experiences, Lessons and Future Directions* (pp. 93–109).

Butler, R. a, & Laurance, W. F. (2009). Is oil palm the next emerging threat to the Amazon? *Tropical Conservation Science, 2,* 1–10. doi:papers2://publication/uuid/5F4271A6-E2C2-4771-9248-7C5BC701B38B

Cabo, F. S., Trajanoski, Z., & Cho, K. (2003). A Graphical Users Interface to Normalize Microarray Data. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)* (pp. 1–16).

Carcassoni, M., & Hancock, E. R. (2000). An improved point proximity matrix for modal matching. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000, 2,* 34–37. doi:10.1109/ICPR.2000.906013

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., & Franklin, J. (1994). A Neural Network Method for Efficient Vegetation Mapping, *4257*(99).

Chandwani, V., Agrawal, V., & Nagar, R. (2015). Modeling slump of ready mix concrete using genetic algorithms assisted training of Artificial Neural Networks. *Expert Systems with Applications, 42*(2), 885–893. doi:10.1016/j.eswa.2014.08.048

Cho, J. H., Kim, Y. W., Na, K. J., & Jeon, G. J. (2008). Wireless electronic nose system for real-time quantitative analysis of gas mixtures using micro-gas sensor array and neuro-fuzzy network. *Sensors and Actuators, B: Chemical, 134,* 104–111. doi:10.1016/j.snb.2008.04.019

Clancey, W. J. (1985). Heuristic Classification, (June).

Cohen, U., Allison, D., & Witte, J. (2009). *Critical Issues in Healthcare Environments* (pp. 1–71).

Cordeiro, J. R., Martinez, M. I. V., Li, R. W. C., Cardoso, A. P., Nunes, L. C., Krug, F. J., … Gruber, J. (2012). Identification of Four Wood Species by an Electronic Nose and by LIBS. *International Journal of Electrochemistry, 2012,* 1–5. doi:10.1155/2012/563939

Cornillon, P. A., Guyader, A., Husson, F., Jegou, N., Josse, J., Kloareg, M., … Rouvière, L. (2012). *R for Statistics.* Taylor & Francis. Retrieved from https://books.google.com.my/books?id=KbV_y3at7ZkC

Craven, M. A. (1995). Electronic noses - development and future prospects, *15*(9).

D.L.Hartmann. (2013). Matrix Methods for Analysis of Structure in Data Sets. In *Matrix Methods* (pp. 61–103).

D.Stone, & Ellis, J. (n.d.). *Calibration and Linear Regression Analysis: A Self-Guided Tutorial.*

Dean, S., Illowsky, B., & Ph, D. (2009). Descriptive Statistics : Box Plot This work is produced by The Connexions Project and licensed under the Creative Commons Attribution License †, 2–5.

DeJordy, R., Borgatti, S. P., Roussin, C., & Halgin, D. S. (2007). Visualizing Proximity Data. *Field Methods, 19*(3), 239–263. doi:10.1177/1525822X07302104

Department of Environment. (2010). Environmental Requirements : A Guide For Investors.

DeSilva, F., & Gorrell, M. (2000). *Wastewater Technology Fact Sheet Ammonia Stripping.*

Devaney, M., & Cheetham, B. (2005). Case-Based Reasoning for Gas Turbine Diagnostics.

Dobrev, D. (2004). A Definition of Artificial Intelligence, 1–7.

Drake, P. (n.d.). Regression step-by-step using Microsoft Excel ®.

Drew, C. H., Grace, D. A., Silbernagel, S. M., Hemmings, E. S., Smith, A., Griffith, W. C., … Faustman, E. M. (2003). Nuclear waste transportation: Case studies of identifying stakeholder risk information needs. *Environmental Health Perspectives.* doi:10.1289/ehp.5203

El-gawad, H. A., & Ramzy, E. M. (2013). Purification and characterization of toxic waste in the aquatic environment using common carp , Cyprinus carpio. *Journal of Natural Resources and Development, 3*, 27–34. doi:10.5027/jnrd.v3i0.03

EMSURE. (2014). Ammonia solution 28-30 %.

Esty, W. W., & Banfield, J. D. (2003). The Box-Percentile Plot. *Journal Of Statistical Software, 8*, 1–14. Retrieved from http://www.jstatsoft.org/v08/i17/paper

Fischer, M. M., & Gopal, S. (1996). Network Models : Backpropagation in a Spatial Interaction Modeling Context, *1*(1).

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: statistical learning of visual feature hierarchies. *Journal of Experimental Psychology. General, 134*(4), 521–37. doi:10.1037/0096-3445.134.4.521

FitzGerald, J., & Tol, R. S. J. (2007). Airline emissions of carbon dioxide in the European trading system. *CESifo Forum, 8*, 51–54.

Floyd, M. W., Esfandiari, B., & Lam, K. (2008). A Case-based Reasoning Approach to Imitating RoboCup Players, 251–256.

Gan, G., Ma, C., & Wu, J. (2007). *Data Clustering: Theory, Algorithms, and Applications.*

Gao, D., Liu, F., & Wang, J. (2012). Quantitative analysis of multiple kinds of volatile organic compounds using hierarchical models with an electronic nose. *Sensors and Actuators, B: Chemical, 161,* 578–586. doi:10.1016/j.snb.2011.11.003

García-Diego, F.-J., Sánchez-Quinche, A., Merello, P., Beltrán, P., & Peris, C. (2013). Array of Hall Effect sensors for linear positioning of a magnet independently of its strength variation. a case study: monitoring milk yield during milking in goats. *Sensors (Basel, Switzerland), 13*(6), 8000–12. doi:10.3390/s130608000

Gayer, G., Gilboa, I., & Lieberman, O. (2007). Rule-Based and Case-Based Reasoning in Housing Prices. *Journal of Theoretical Economics, 7*(1).

Gomez, G. (2010). Normalization methods and data preprocessing.

Granta Y Nakayama. (2007). US Environmental Protection Agency Office of Enforcement and Compliance. *Water.* Retrieved from www.epa.gov/compliance/resources/policies/ej/ej-testimony-7-25-07.pdf

Hapse, S. A., Kadaskar, P. T., & Shirsath, A. S. (2011). Difference spectrophotometric estimation and validation of ibuprofen from bulk and tablet dosage form, *3*(6), 18–23.

Harvey, D. (200AD). *Modern Analytical Chemistry.* (K. T. Kane & J. L. Bensink, Eds.) (first., p. 30). James M. Smith.

Haykin, S. (1998). Neural Networks : *Convergence,* 1–16. doi:10.1109/79.487040

Henriet, J., Leni, P.-E., Laurent, R., & Salomon, M. (2014). Case-Based Reasoning adaptation of numerical representations of human organs by interpolation. *Expert Systems with Applications, 41*(2), 260–266. doi:10.1016/j.eswa.2013.05.064

Hodgins, D. (1995). The development of an electronic ' nose ' for industrial and environmental applications. *Sensors and Actuators B: Chemical, 26-27,* 255–258.

Huang, S.-C., & Yeh, C.-H. (2013). Image contrast enhancement for preserving mean brightness without losing image features. *Engineering Applications of Artificial Intelligence, 26*(5-6), 1487–1492. doi:10.1016/j.engappai.2012.11.011

Hunt, K. (2010). Box Plots. In *Discovering Advanced Algebra Condensed Lessons* (pp. 17–22). Retrieved from http://math.kendallhunt.com/x3101.html

Iglesias, F., & Kastner, W. (2013). Analysis of Similarity Measures in Times Series Clustering for the Discovery of Building Energy Patterns. *Energies*, *6*(2), 579–597. doi:10.3390/en6020579

Inaoka, M. (2005). Environmental Pollution Control Measures. In M. KOMASAWA, K. ADACHI, & Y. DOGUCHI (Eds.), *Japan's Experience in Public Health and Medical Systems* (pp. 145–164).

Intcal, E. T. H. E., Methodologies, S., Heaton, M. N. T. J., & Buck, P. G. B. C. E. (2013). The bayesian approach to radiocarbon calibration curve estimation: the intcal13, marine13, and shcal13 methodologies, *55*(4), 1905–1922.

Isotalo, J. (n.d.). *Basics of Statistics* (pp. 0–82).

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*, 4–37. doi:10.1109/34.824819

Jha, G. K. (n.d.). Artificial neural networks.

Jonassen, D. H., & Hernandez-serrano, J. (2002). Case-Based Reasoning and Instructional Design : Using Stories to Support Problem Solving. *ETR&D*, *50*(2), 65–77.

Jørgensen, M. W., & Phillips, L. J. (2002). *Discourse Analysis as Theory and Method* (p. 229). SAGE Publications. Retrieved from https://books.google.com/books?id=ILdm1r1IsfMC&pgis=1

Juszczak, P. (n.d.). Feature scaling in support vector data description.

Kampstra, P. (2008). Beanplot : A Boxplot Alternative for Visual Comparison of Distributions. *Journal Of Statistical Software*, *28*(November), 1–9. Retrieved from http://www.jstatsoft.org/v28/c01/paper

Kim, K. H. (2010). Experimental demonstration of masking phenomena between competing odorants via an air dilution sensory test. *Sensors*, *10*, 7287–7302. doi:10.3390/s100807287

Kim, K. H. (2011). The averaging effect of odorant mixing as determined by air dilution sensory tests: A case study on reduced sulfur compounds. *Sensors*, *11*, 1405–1417. doi:10.3390/s110201405

Kim, K. H., & Park, S. Y. (2008). A comparative analysis of malodor samples between direct (olfactometry) and indirect (instrumental) methods. *Atmospheric Environment*, *42*, 5061–5070. doi:10.1016/j.atmosenv.2008.02.017

Klotz-Ingram, C., & Day-Rubenstein, K. (1999). The changing agricultural research
environment: What does it mean for public-private innovation? *AgBioForum*.

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence
Review, 6*(1), 3–34. Retrieved from
http://www.springerlink.com/index/10.1007/BF00155578

Kolodner, J. L., & Leake, D. B. (1996). A Tutorial Introduction to Case-Based
Reasoning. In *Case-Based Reasoning: Experiences, Lessons and Future Directions*
(pp. 31–65).

Koroleff, F. (1966). Direct spectrophotometric determination of ammonia in
precipitation. *Tellus XVIII, 2*, 562–565.

LabChem. (2013). Ammonium Hydroxide , 28-30 % w/w Safety Data Sheet.

Lee, J. (2003). Introduction to software engineering with computational intelligence.
Information and Software Technology, 45(7), 371–372. doi:10.1016/S0950-
5849(03)00009-0

Leydesdorff, L. (n.d.). On the Normalization and Visualization of Author Co-Citation
Data : Salton ' s Cosine versus the Jaccard Index. *Journal of the American Society of
Information Science & Technology, 131*.

Leydesdorff, L., & Vaughan, L. (2006). Co-occurrence matrices and their applications in
information science: Extending ACA to the Web environment. *Journal of the
American Society for Information Science and Technology, 57*(12), 1616–1628.
doi:10.1002/asi.20335

Liu, G., Yang, Z., Chen, B., Zhang, Y., Su, M., & Zhang, L. (2013). Emergy evaluation
of the urban solid waste handling in liaoning province, China. *Energies, 6*, 5486–
5506. doi:10.3390/en6105486

Loan, D. K., Con, T. H., Hong, T. T., Thi, L., & Ly, M. (2013). Quick Determination of
Ammonia Ions in Water Environment Based on Thymol Color Creating Reaction.
Environmental Sciences, 1(2), 83–92.

Lyman, R. O., & Michael Longnecker. (2010). *An Introduction to Statistical Methods
and Data Analysis*. (M. Taylor, Ed.) (Sixth Edit.).

Magnier, L., & Haghighat, F. (2010). Multiobjective optimization of building design
using TRNSYS simulations, genetic algorithm, and Artificial Neural Network.
Building and Environment, 45(3), 739–746. doi:10.1016/j.buildenv.2009.08.016

Mariano, J., & La Rovere, E. (2007). Oil and gas exploration and production activities in
Brazil: The consideration of environmental issues in the bidding rounds promoted

by the National Petroleum Agency. *Energy Policy, 35*(5), 2899–2911. doi:10.1016/j.enpol.2006.10.005

Marinković, N., Vitale, K., Afrić, I., & Janev Holcer, N. (2005). Hazardous medical waste management as a public health issue. *Arhiv Za Higijenu Rada I Toksikologiju, 56*, 21–32.

Marmolejo-ramos, F., & Tian, T. S. (2010). The shifting boxplot . A boxplot based on essential summary statistics around the mean . *International Journal of Psychological Research, 3*(1), 37–45.

Marsh, K., & Bugusu, B. (2007). Food packaging--roles, materials, and environmental issues. *Journal of Food Science, 72*(3), R39–55. doi:10.1111/j.1750-3841.2007.00301.x

Martin, B., & Hach, C. C. (2002). An Introduction to Standards and Quality Control for the Laboratory In memory of, (2426).

Mažeikienė, A., & Dauknys, R. (2014). Removal of ammonium ions from digested sludge liquors using natural sorbent zeolite. *The 9th International Conference "Environmental Engineering 2014,"* (May). doi:10.3846/enviro.2014.087

Men, H., Liu, H., Pan, Y., Wang, L., & Zhang, H. (2011). Electronic Nose Based on an Optimized Competition Neural Network. *Sensors*. doi:10.3390/s110505005

Miche, Y., Sorjamaa, A., Bas, P., Simula, O., Jutten, C., & Lendasse, A. (2010). OP-ELM: optimally pruned extreme learning machine. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council, 21*(1), 158–62. doi:10.1109/TNN.2009.2036259

Miecznikowski, J. C., & Sellers, K. F. (2002). Statistical Analysis of Chemical Sensor Data. In *Advances in Chemical Sensors* (pp. 327–347).

Miles, J., & Philip Banyard. (2007). *Understanding and Using Statistics in Psychology: A Practical Introduction* (pp. 61–75).

Mishra, A. K., Kumar, M., Mishra, A., & Verma, A. (2010). Validated UV spectroscopic method for estimation of Salbutamol from tablet formulations, *2*(3), 207–211.

Mohamed, N. S., Zainudin, S., & Othman, Z. A. (2014). Classification Techniques for Drug Response Microarray Data. *Journal of Next Generation Information Technology, 5*(3), 2014.

Mok, P. Y., Huang, H. Q., Kwok, Y. L., & Au, J. S. (2012). A robust adaptive clustering analysis method for automatic identification of clusters. *Pattern Recognition, 45*(8), 3017–3033. doi:10.1016/j.patcog.2012.02.003

Najib, M. S., Ahmad, M. U., Funk, P., Taib, M. N., & Ali, N. A. M. (2012). Agarwood
Classification : A Case-based Reasoning Approach Based on E-nose. In *IEEE 8th
International Colloqium on Signal Processing and Its Application* (pp. 120–126).

Najib, M. S., Taib, M. N., Ali, N. A. M., Arip, M. N. M., & Jalil, A. M. (2011).
Classification of Agarwood grades using ANN. *International Conference on
Electrical, Control and Computer Engineering 2011 (InECCE)*, 367–372.
doi:10.1109/INECCE.2011.5953908

Nakkrasae, S., & Edwards, W. R. (n.d.). Fuzzy Subtractive Clustering Based Indexing
Approach for Software Components Classification.

Natsoulis, G., Ghaoui, L. El, Lanckriet, G. R. G., Tolley, A. M., Leroy, F., Dunlea, S., ...
Jarnagin, K. (2005). Classification of a large microarray data set : Algorithm
comparison and analysis of drug signatures, 724–736.
doi:10.1101/gr.2807605.Deriving

Nau, R. (2014). Review of basic statistics and the simplest forecasting model: the sample
mean, (August), 1–18.

Navidi, W. (2014). *Statistics for Engineers and Scientists.* McGraw-Hill Education.
Retrieved from https://books.google.com.my/books?id=ck5gnAEACAAJ

Negnevitsky, M. (2005). *Artificial Intelligence: A Guide to Intelligent Systems* (Second.,
p. 415). Retrieved from http://www.amazon.com/Artificial-Intelligence-Intelligent-
Systems-Edition/dp/0321204662

Nicolas, J., Romain, a.-C., Wiertz, V., Maternova, J., & André, P. (2000). Using the
classification model of an electronic nose to assign unknown malodours to
environmental sources and to monitor them continuously. *Sensors and Actuators B:
Chemical*, 69(3), 366–371. doi:10.1016/S0925-4005(00)00487-1

Ning, R., Li, Q., & Xu, S. (2015). Research on the Fitting Methods of Calibration Curve
for Atomic Emission Spectrum Analysis, (Iiicec), 1047–1052.

Noble, S., Neville, D., & Houghton, R. (2014). Determination of 8-iso-prostaglandin
F(2α) (8-iso-PGF(2α)) in human urine by ultra-performance liquid chromatography-
tandem mass spectrometry. *Journal of Chromatography. B, Analytical Technologies
in the Biomedical and Life Sciences*, 947-948, 173–8.
doi:10.1016/j.jchromb.2013.12.028

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., & Hayasaka, T.
(2007). An Asian emission inventory of anthropogenic emission sources for the
period 1980–2020. *Atmospheric Chemistry and Physics Discussions*.
doi:10.5194/acpd-7-6843-2007

Ontanon, S., & Ram, A. (2011). Case-Based Reasoning and User-Generated AI for Real-Time Strategy Games, *30332*, 1–22.

Orlov, M. L. (1996). *MULTIPLE LINEAR REGRESSION ANALYSIS USING MICROSOFT EXCEL*.

Orubu, C., Odusola, A., & Ehwarieme, W. (2004). The Nigerian oil industry: environmental diseconomies, management strategies and the need for community involvement. *Journal of Human Ecology, 16*, 203–214. Retrieved from http://www.krepublishers.com/02-Journals/JHE/JHE-16-0-000-000-2004-Web/JHE-16-3-151-226-2004-Abst-PDF/JHE-16-3-203-214-2004-1168-Orubu-C-O/JHE-16-3-203-214-2004-Orubu-C-O.htm

Ozkan, M., Dawant, B. M., & Maciunas, R. J. (1993). Neural-Network-B ased Segmentation of Multi-Modal Medical Images : A Comparative and Prospective Study, *12*(3), 534–544.

Pace, L. (2012). Correlation and Regression. In *Beginning R: An Introduction to Statistical Programming* (pp. 215–242). Apress.

Pal, S. K., & Shiu, S. C. K. (2004). *Foundations of Soft Case-Based Reasoning. Concepts Features and Soft Computing Applied In* (p. 274). doi:10.1023/B:APIN.0000043556.29968.81

Pedrycz, W., & Waletzky, J. (1997). Fuzzy Clustering in Software Reusability. *Software-Practice and Experience, 27*(MARCH), 245–270.

Perales, F. J., Campilho, A. J. C., Blanca, N. P. de la, & Sanfeliu, A. (2003). *Pattern Recognition and Image Analysis* (p. 1142). Springer Science & Business Media. Retrieved from https://books.google.com/books?id=dpfo7UxBQi0C&pgis=1

Pfannkuch, M. (2006). Comparing Box Plot Distributions: A Teacher's Reasoning. *Statistics Education Research Journal, 5*(2), 27–45.

Potter, K. (n.d.). Methods for Presenting Statistical Information : The Box Plot a b c, 97–106.

R.R. Frerichs. (2008). Simple Random Sampling. In *Rapid Surveys*.

Rabbi, A., Nasouri, K., Bahrambeygi, H., Shoushtari, A. M., & Babaei, M. R. (2012). RSM and ANN approaches for modeling and optimizing of electrospun polyurethane nanofibers morphology. *Fibers and Polymers, 13*(8), 1007–1014. doi:10.1007/s12221-012-1007-x

Rahman, R. O. A., Ibrahium, H. A., & Hung, Y.-T. (2011). Liquid Radioactive Wastes Treatment: A Review. *Water*. doi:10.3390/w3020551

Ramamoorthy, H. V., Mohamed, S. N., & Devi, D. S. (2014). E-Nose and E-Tongue: Applications and Advances in Sensor Technology. *Journal of NanoScience and NanoTechnology*. Retrieved from http://indiasciencetech.com/index.php?journal=nanotechnology&page=article&op=view&path[]=71

Ramos-Pollan, R., Guevara-Lopez, M. A., & Oliveira, E. (n.d.). Introducing ROC curves as error measure functions . A new approach to train ANN-based biomedical data, (1).

Ramsey, C. B., Scott, E. M., & Plicht, J. van der. (2013). Calibration for archaeological and environmental terrestrial samples in the time range 26 – 50 ka cal bp. *Radiocarbon*, *55*(4), 2021–2027.

Randall, N. P., & James, K. L. (2012). The effectiveness of integrated farm management, organic farming and agri-environment schemes for conserving biodiversity in temperate Europe - A systematic map. *Environmental Evidence*. doi:10.1186/2047-2382-1-4

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., ... Plicht, J. van der. (2013). Intcal13 and Marine13 Radiocarbon Age Calibration Curves 0 – 50,000 years cal bp. *Radiocarbon*, *55*(4), 1869–1887.

Reiner, P. D. (2015). *Algorithms for Optimal Construction and Training of Radial Basis Function Neural Networks*.

Rezakazemi, M., Razavi, S., Mohammadi, T., & Ghafari, A. (2011). Simulation and determination of optimum conditions of pervaporative dehydration of isopropanol process using synthesized PVA – APTEOS / TEOS nanocomposite membranes by means of expert systems. *Journal of Membrane Science*, *379*(1-2), 224–232. doi:10.1016/j.memsci.2011.05.070

Richter, M. M., & Aamodt, A. (2005). Case-based reasoning foundations. *The Knowledge Engineering Review*. doi:10.1017/S0269888906000695

Rodriguez, H. M., & J.O, A. (2013). Simple Boxplot. In J. P. Donnelly (Ed.), *Building SPSS Graphs to Understand Data* (pp. 200–209).

Roopadevi, H., & Somashekar, R. K. (2012). Assessment of the toxicity of waste water from a textile industry to Cyprinus carpio. *Journal of Environmental Biology*, *33*, 167–171.

Rousseeuw, P. J., Ruts, I., & Tukey, J. W. (2012). The Bagplot : A Bivariate Boxplot, (October), 37–41.

Sanders, D. H., & Smidt, R. K. (2000). *Statistics: A First Course*. McGraw-Hill. Retrieved from https://books.google.com.my/books?id=wF7cAAAACAAJ

Sarkar, K., Ghalia, M. Ben, Wu, Z., & Bose, S. C. (2009). A neural network model for the numerical prediction of the diameter of electro-spun polyethylene oxide nanofibers. *Journal of Materials Processing Technology, 209*(7), 3156–3165. doi:10.1016/j.jmatprotec.2008.07.032

Schau, C., & EMMİOĞLU, E. (2012). DO INTRODUCTORY STATISTICS COURSES IN THE UNITED STATES IMPROVE STUDENTS ' ATTITUDES ? 7. *Statistics Education Research Journal, 11*(2), 86–94.

Schiller, F., Jena, D.-, & Curriet, L. A. (1998). INTERNATIONAL UNION OF PURE COMMISSION ON GENERAL ASPECTS OF ANALYTICAL CHEMISTRY * GUIDELINES FOR CALIBRATION IN ANALYTICAL CHEMISTRY Guideline for calibration in analytical chemistry- Part 1 . Fundamentals and single component, *70*(4), 993–1014.

Schwarz, C. J. (2013). Correlation and simple linear regression., *404*, 876–941.

Scott, G. L., & Longuet-Higgins, H. C. (1990). Feature grouping by "relocalisation" of eigenvectors of the proximity matrix. *Procedings of the British Machine Vision Conference 1990*, 20.1–20.6. doi:10.5244/C.4.20

Scott, L. (n.d.). Regression Analysis Tutorial.

Seoane, J. A., Day, I. N. M., Campbell, C., Casas, J. P., & Gaunt, T. R. (n.d.). Using a Random Forest proximity measure for variable importance stratification in genotypic data, 1049–1060.

Shakhashiri, P. (2008). Ammonia, NH3, *3*, 2.

Sharif, M. S., Abbod, M., Amira, A., & Zaidi, H. (2010). Artificial neural network-based system for PET volume segmentation. *International Journal of Biomedical Imaging, 2010*. doi:10.1155/2010/105610

Shiavi, R. (2010). *Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications* (p. 424). Academic Press. Retrieved from https://books.google.com/books?id=1qGPVgjnx4gC&pgis=1

Shonnard, D. R. (n.d.). Chapter 1. An Introduction to Environmental Issues by David R. Shonnard Environmental issues gained increasing prominence in the latter half of the 20, 1–2.

Sigma -Aldrich. (2003). Ammonium Hydroxide.

Silva, R. de C. F. S., Almeida, D. G., Rufino, R. D., Luna, J. M., Santos, V. A., & Sarubbo, L. A. (2014). Applications of biosurfactants in the petroleum industry and the remediation of oil spills. *International Journal of Molecular Sciences.* doi:10.3390/ijms150712523

Simonoff, J. S. (2008). Statistical analysis using Microsoft Excel Microsoft, pp. 1–12.

Sohn, J. H., Smith, R., Yoong, E., Leis, J., & Galvin, G. (2003). Quantification of Odours from Piggery Effluent Ponds using an Electronic Nose and an Artificial Neural Network. *Biosystems Engineering, 86*, 399–410. doi:10.1016/j.biosystemseng.2003.09.003

Song, K., Wang, Q., Liu, Q., Zhang, H., & Cheng, Y. (2011). A Wireless Electronic Nose System Using a Fe2O3 Gas Sensing Array and Least Squares Support Vector Regression. *Sensors.* doi:10.3390/s110100485

Status, T., & Disease, N. (1999). Sensitivity and Specificity, 1–3.

Stolcke, A., Kajarekar, S., Ferrer, L., International, S. R. I., & Park, M. (2008). Nonparametric Feature Normalization for SVM-Based Speaker Verification, 1577–1580.

Suah, F. B. M., Ahmad, M., & Taib, M. N. (2003). Optimisation of the range of an optical fibre pH sensor using feed-forward artificial neural network. *Sensors and Actuators B: Chemical, 90*(1-3), 175–181. doi:10.1016/S0925-4005(03)00025-X

Subasi, A., & Erçelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine, 78*(2), 87–99. doi:10.1016/j.cmpb.2004.10.009

Sun, Y., & Genton, M. G. (2011). Functional Boxplots. *Journal of Computational and Graphical Statistics, 20*(2), 316–334. doi:10.1198/jcgs.2011.09224

Szczurek, A., & Maciejewska, M. (2013). The Concept of System for the Measurement of Volatile Organic Compounds. *Sensors and Actuators B: Chemical, 35*(1), 39–46.

Taghizadeh, S. R. (2000). Digital Signal Processing, (January), 1–48.

Taib, M. N., Andres, R., & Narayanaswamy, R. (1996). Extending the response range of an optical fibre pH sensor using an artificial neural network, *2670*(96).

Taib, M. N., & Narayanaswamy, R. (1997). Multichannel calibration technique for optical-fibre chemical sensor using artificial neural network, *39*, 365–370.

Triska, R. (n.d.). Artificial Intelligence , classification theory and the uncertainty reduction process, 479–483.

Tudu, B., Kow, B., Bhattacharyya, N., & Bandyopadhyay, R. (2009). Normalization techniques for gas sensor array as applied to classification for black tea. *International Journal on Smart Sensing and Intelligent Systems, 2*(1), 176–189.

U.S Environmental Protection Agency. (2013). Total Nitrogen. Retrieved January 07, 2015, from www.epa.gov/owow/monitoring/volunteer/stream

Umpierre, D., Ribeiro, P. a B., Schaan, B. D., & Ribeiro, J. P. (2013). Volume of supervised exercise training impacts glycaemic control in patients with type 2 diabetes: a systematic review with meta-regression analysis. *Diabetologia, 56*(2), 242–51. doi:10.1007/s00125-012-2774-z

UNEP Global Environmental Alert Service (GEAS). (2011), (December 2010), 1–8.

Van Loco, J., Elskens, M., Croux, C., & Beernaert, H. (2002). Linearity of calibration curves: use and misuse of the correlation coefficient. *Accreditation and Quality Assurance, 7*(7), 281–285. doi:10.1007/s00769-002-0487-6

Veloso, M. M. (1996). Flexible Strategy Learning: Analogical Replay of Problem Solving Episodes. In *Case-Based Reasoning: Experiences, Lessons and Future Directions* (pp. 137–149).

Venet, D. (2003). MatArray: a Matlab toolbox for microarray data. *Bioinformatics, 19*(5), 659–660. doi:10.1093/bioinformatics/btg046

Vergara, A., & Llobet, E. (2012). Sensor Selection and Chemo-Sensory Optimization: Toward an Adaptable Chemo-Sensory System. *Frontiers in Neuroengineering.* doi:10.3389/fneng.2011.00019

Voskoglou, M. G. (2009). Fuzzy Sets in Case-Based Reasoning. *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 6*, 345–373. doi:10.1109/FSKD.2009.667

Vries, H. De. (1993). THE ROWWISE CORRELATION BETWEEN TWO PROXIMITY MATRICES AND THE PARTIAL ROWWISE CORRELATION. *Psychometrika, 58*(1), 53–69.

Wickham, H., & Stryjewski, L. (2011). 40 years of boxplots, 1–17.

Williamson, D. F., D, P., Parker, R. A., & Kendrick, J. S. (1989). The Box Plot : A Simple Visual Method to Interpret Data, *110*(11), 916–921.

Wilson, K. A., Green, N. D., Agrawal, L., Madhusoodanan, D., Riley, B., & Sigmon, J. P. (n.d.). Graph-based Proximity Measures (pp. 135– 166).

Wilson, T. A. (2007). The Total and Partial Vapor Pressures of Aqueous Ammonia Solutions. *The Engineering Experiment Station.*

Winter, J. B. A., Ballinger, D. G., Berg, E. L., Booth, R. L., Bordner, R. H., Britton, P. W., ... Winter, J. A. (2001). *Handbook for Analytical Quality Control in Wastewater Laboratories.*

Wochnowski, C., Metev, S., & Sepold, G. (2000). UV–laser-assisted modification of the optical properties of polymethylmethacrylate. *Applied Surface Science, 154-155,* 706–711. doi:10.1016/S0169-4332(99)00435-3

Wong, H. B., & Lim, G. H. (2011). Measures of Diagnostic Accuracy : Sensitivity , Specificity , PPV and NPV, *20*(4), 316–318.

Wu, W., May, R., Dandy, G. C., & Maier, H. R. (2012). A method for comparing data splitting approaches for developing hydrological ANN models. *International Environmental Modelling and Software Society.*

Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks.* doi:10.1109/TNN.2005.845141

Yetilmezsoy, K., & Sapci-Zengin, Z. (2009). Recovery of ammonium nitrogen from the effluent of UASB treating poultry manure wastewater by MAP precipitation as a slow release fertilizer. *Journal of Hazardous Materials, 166*(1), 260–9. doi:10.1016/j.jhazmat.2008.11.025

Yoshio, A. (n.d.). The Formation of Heavy Industry. *The Formation of Heavy Industry.*

Yosinski, J., & Paffenroth, R. (2000). *Nonlinear Estimation for Arrays of Chemical Sensors.*

Youn, E., & Jeong, M. K. (2009). Class dependent feature scaling method using naive Bayes classifier for text datamining. *Pattern Recognition Letters, 30*(5), 477–485. doi:10.1016/j.patrec.2008.11.013

Zellers, E. T., Pan, T.-S., Patrash, S. J., Han, M., & Batterman, S. a. (1993). Extended disjoint principal-components regression analysis of SAW vapor sensor-array responses. *Sensors and Actuators B: Chemical, 12*(2), 123–133. doi:10.1016/0925-4005(93)80008-Y

Zhu, W., Zeng, N., & Wang, N. (2010). NESUG 2010 Health Care and Life Sciences Sensitivity , Specificity , Accuracy , Associated Confidence Interval and ROC Analysis with Practical SAS ® Implementations K & L consulting services , Inc , Fort Washington , PA Octagon Research Solutions , Wayne , 1–9.

APPENDIX A

List of publication as author

Zahari, M. F., Julius, T. A., Halim, F. A., Najib, M. S., Ghazali, K. H., & Mohd Azoddein, A. A. (2015). Intelligent Classification Hazardous Gas Using Sensors Array. Journal of Advances in Information Technology, Volume 6 (No. 4), 233–237.

Zahari, M. F., Julius, T. A., Halim, F. A., Najib, M. S., Ghazali, K. H., & Mohd Azoddein, A. A. (2015). Intelligent Classification Hazardous Gas using Sensors Array. Presented at The 8th International Conference on Advanced Computer Theory and Engineering.

Zahari, M. F., Najib, M. S., Ghazali, K. H., Halim, F. A., & Mohd Azoddein, A. A. (2014). Classification of Ammonia in water for Oil and Gas Industry using Case Based Reasoning (CBR). Presented at Colloquium on Robotics, Unmanned Systems And Cybernetics 2014 (CRUSC 2014) (Vol.2014).

Zahari, M. F., Najib, M. S., Ghazali, K. H., Halim, F. A., Tajuddin, S. N., & Latib, E. H. A. (2014). Gaharu Sensor : Classification Using Case Based Reasoning (CBR). Presented at Colloquium On Robotics, Unmanned Systems And Cybernetics 2014 (CRUSC 2014) (Vol. 2014).

List of publication as co-author

Halim, F. A., Najib, M. S., Ghazali, K. H., Mohd Azoddein, A. A. & Zahari, M. F., Classification of Ammonia Odor-profile in Fertilizer. International Journal of Advances in Science, Engineering and Technology (IJASEAT), Volume-3, Issue-4, Spl,-3, pp 1-5, 2015.

Halim, F. A., Najib, M. S., Ghazali, K. H., & Zahari, M. F. (2014). Classification of Ammonia Odor-profile Using k-NN Technique. Presented at Colloquium On Robotics, Unmanned Systems And Cybernetics 2014 (CRUSC 2014) (Vol. 2014, pp. 4–8).

Halim, F. A., Najib, M. S., Ghazali, K. H., Mohd Azoddein, A. A. & Zahari, M. F.., 2015. "Classification of Ammonia Odor-profile in Fertilizer," Presented at 2nd International Conference on Modern Trends in Engineering and Technology (ICMTET)