

# Prediction of Quality for Different Type of Wine based on Different Feature Sets Using Supervised Machine Learning Techniques

Satyabrata Aich<sup>\*</sup>, Ahmed Abdulhakim Al-Absi<sup>\*\*</sup>, Kueh Lee Hui<sup>\*\*\*</sup>, and Mangal Sain<sup>\*\*\*\*</sup>

<sup>\*</sup>Department of Computer Engineering, Inje University, South Korea

<sup>\*\*</sup>Department of Computer Engineering, Kyungdong University- Global Campus, Gangwondo, South Korea

<sup>\*\*\*</sup>Dept. of Electrical Engineering, Dong-A University, South Korea

<sup>\*\*\*\*</sup>Department of Computer Engineering, Dongseo University, South Korea

satyabrataaich@gmail.com, absiahmed@kduniv.ac.kr, leehkueh@dau.ac.kr, mangalsain1@gmail.com

**Abstract—** In recent years, most of the industries promoting their products based on the quality certification they received on the products. The traditional way of assessing the product quality is time consuming, however with the invent of machine learning techniques the processes has become more efficient and consumed less time than before. In this paper we have explored, some of the machine learning techniques to assess the quality of wine based on the attributes of wine that depends on quality. We have used white wine and red wine quality dataset for this research work. We have used different feature selection technique such as genetic algorithm (GA) based feature selection and simulated annealing (SA) based feature selection to check the prediction performance. We have used different performance measure such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value for comparison using different feature sets and different supervised machine learning techniques. We have used nonlinear, linear and probabilistic classifiers. We have found that feature selection-based feature sets able to provide better prediction than considering all the features for performance prediction. We have found accuracy ranging from 95.23% to 98.81% with different feature sets. This analysis will help the industries to access the quality of the products at less time and more efficient way.

**(Keywords—** Machine learning; feature selection; classifiers; performance metrics; wine quality



**Satyabrata Aich** is working as a researcher in the field of computer engineering. He has over four years of teaching, research and industry experience in India and abroad. He has published many research papers in journals and conferences in the realms of Supply Chain Management and data analytics. His research interests are natural language processing, Machine learning, supply chain management, data mining.



**Kueh Lee Hui** is working as an assistant professor at the department of Electrical Engineering, Dong-A University since 2012. She completed her PhD Degrees from Department of Electrical Engineering, Dong-A University, Korea. In 2009 she completed her BS degree in Electronic and Communication, Department of Electronic Engineering, University Malaysia of Sarawak, Malaysia. She also done MS in 2007 from Malaysia. Her research interests are image processing, face recognition, digital image forensic, intelligent control and control application, power system.



**Ahmed Abdulhakim Al-Absi** is an assistant professor in Department of Computer Engineering (Smart Computing) at Kyungdong University in South Korea. He earned a Ph.D. in computer science from Dongseo University in 2015. He received M.Sc. degree in information technology at University Utara Malaysia in 2011, and B.Sc. degree in computer applications at Bangalore University in 2008. His research interests include Big Data processing, Hadoop, Cloud computing, IoT, Distributed systems, Parallel computing, Bioinformatics, Security, and VANETs.



**Mangal Sain** received the M.Sc. degree in computer application from India in 2003 and the Ph.D. degree in computer science in 2011. Since 2012, he has been an Assistant Professor with the Department of Computer Engineering, Dongseo University, South Korea. His research interest includes wireless sensor network, cloud computing, Internet of Things, embedded systems, and middleware. He has authored over 50 international publications including journals and international conferences. He is a member of TIIS and a TPC member of more than ten international conferences.