

# **Supporting consumer decisions in cloud computing using trust model**

**by Muhammad Raza**

Thesis submitted in fulfilment of the requirements for  
the degree of

**Doctor of Philosophy**

under the supervision of Associate Professor Farookh  
Khadeer Hussain

University of Technology Sydney  
Faculty of Engineering and Information Technology

May 2021

# CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Muhammad Raza declare that this thesis, is submitted in fulfilment of the requirements for the award of PhD, in the School of Computer Science, Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:  
Signature: Signature removed prior to publication.  
Date: 26-04-2021

# Acknowledgements

I would like to express my sincere gratitude to my principal supervisor A/Prof. Farookh Khadeer Hussain for his continuous support, guidance and encouragement during my Ph.D studies. His generosity and patience always provided me with extra energy throughout difficult times. I would also like to thank my co-supervisor A/ Prof. Zia ur Rehman for his help and support. His presence and cooperation allowed me to continue and complete my Ph.D journey.

I specially thank A/Prof. Omar Khadeer Hussain for his guidance, support and untiring willingness to help that kept me continue my research journey. I would also like to express my appreciation to Mr. Ming Zhao for his excellent support and helpful discussions.

I would like to extend my gratitude to my friends at UTS, Dr. Quynh Do, Dr. Supannada Chotipant, Dr. Mohammad Ikram and Dr. Omar Alshaweesh whose company and encouragement during my Ph.D journey have been invaluable.

Finally, I would like to dedicate this thesis to my family. My father, who is not with us to see this accomplishment and my mother, whose passion for reading and learning is the core motivation for my academic choices. I would like to thank my brother and sister for being with me throughout the difficult times and my uncle for his kind support.

# List of Publications

## Journal Articles

1. Raza, M., Hussain, F. K., Hussain, O. K., Zhao, M., & Rehman, Z. u. (2019). A comparative analysis of machine learning models for quality pillar assessment of SaaS services by multi-class text classification of users' reviews. *Future Generation Computer Systems*, 101, 341-371. (*JCR(Q1), Impact Factor (6.125)*)
2. Raza, M., Hussain, F.K., Hussain, O.K. et al. Imputing sentiment intensity for SaaS service quality aspects using T-nearest neighbors with correlation-weighted Euclidean distance. *Knowl Inf Syst* 63, 2541–2584 (2021). *JCR(Q2), Impact Factor (2.936)*)

# Table of Contents

CERTIFICATE OF ORIGINAL AUTHORSHIP .....	ii
Acknowledgements.....	iii
List of Publications .....	iv
Table of Contents.....	v
List of Figures .....	x
List of Tables .....	xiii
Glossary of terms .....	xv
Abstract.....	xviii
<b>Chapter 1: Introduction .....</b>	<b>1</b>
1.1 Introduction .....	1
1.2 Overview of cloud computing .....	1
1.2.1 What is cloud computing? .....	2
1.2.2 Cloud services deployment models .....	3
1.2.3 Cloud services delivery models .....	3
1.3 Key challenges in cloud computing.....	4
1.3.1 Privacy	4
1.3.2 Security	5
1.3.3 Trust	5
1.4 Trust related issues in SaaS .....	6
1.4.1 Factors affecting trust in SaaS .....	6
1.4.2 Missing information regarding trust factors .....	7
1.4.3 Forecasting information regarding trust factors.....	8
1.4.4 Modelling trust for Software as a Service .....	9
1.5 Objectives of the thesis .....	9
1.6 Scope of the thesis .....	9
1.7 Significance of the thesis .....	10
1.8 Plan of the thesis.....	11

1.9	Conclusion .....	12
<b>Chapter 2: Literature Review .....</b>		<b>13</b>
2.1	Introduction .....	13
2.2	Trust assessment and trust models.....	13
2.2.1	Definition of Trust.....	13
2.2.2	Trust modelling approaches.....	14
2.2.3	Trust in different domain .....	15
2.3	Trust assessment and Trust models in Cloud Computing.....	16
2.4	Factors effecting Trust.....	30
2.5	Fuzzy Trust models in cloud computing.....	31
2.6	Critical Analysis .....	32
2.7	Conclusion.....	34
<b>Chapter 3: Problem Definition .....</b>		<b>36</b>
3.1	Introduction .....	36
3.2	Problem definition .....	37
3.3	Research Issues.....	39
3.4	Research Objectives.....	40
3.5	Research approach.....	42
3.6	Conclusion.....	43
<b>Chapter 4: Solution Overview .....</b>		<b>45</b>
4.1	Introduction .....	45
4.2	Trust model for Software as a Service.....	45
4.3	Overview of proposed solution.....	46
4.4	Overview of solution for component 1: Identification of service factors .....	48
4.5	Overview of solution for component 2: Imputing missing sentiment intensity scores for service factors.....	51
4.6	Overview of solution for component 3: Forecasting sentiment intensity scores for service factors.....	53

4.7	Overview of solution for component 4: Modelling trust levels in Software as a Service	55
4.8	Conclusion	57
<b>Chapter 5: An Ensemble approach for identifying SaaS service factors</b>		<b>58</b>
5.1	Introduction	58
5.2	Machine learning based text classification	59
5.3	An Ensemble machine learning approach for multi-class text classification of SaaS reviews	59
5.3.1	Overview of SaaS service factor identification from customer reviews	59
5.3.2	Data preparation	61
5.3.3	Parameter estimation	69
5.3.4	Model evaluation	71
5.3.5	Creating Ensemble	76
5.4	Experiments and Evaluation Results	79
5.4.1	Dataset description	79
5.4.2	Model evaluation	80
5.4.3	Statistical significance test	88
5.4.4	Performance of the Ensemble	88
5.5	Conclusion	90
<b>Chapter 6: T-NN: A Threshold-based Nearest Neighbour approach for Imputing missing values of SaaS service factors</b>		<b>92</b>
6.1	Introduction	92
6.2	Missing data mechanisms	93
6.3	Threshold-based Nearest Neighbour approach	94
6.3.1	Overview of Missing sentiment imputation methodology	94
6.3.2	Data preparation stage	95
6.3.3	Sentiment intensity computation stage	97
6.3.4	Inference stage	98
6.3.5	Validation stage	105
6.4	Experiments and Evaluation results	111
6.4.1	Scheme A	111
6.4.2	Scheme B	116
6.5	Imputing sentiment intensity score of a service factor: An example Case study	129
6.5.1	Finding donor instances	129

6.5.2	Calculating Pearson’s correlation coefficient .....	130
6.5.3	Calculating weighted distance .....	130
6.6	Conclusion .....	132

**Chapter 7: A methodology for time series Forecasting of SaaS service factors 134**

7.1	Introduction .....	134
7.2	Background.....	135
7.2.1	Time series.....	135
7.2.2	Regular and irregular series .....	135
7.2.3	Time series components .....	136
7.2.4	Stationary and non-stationary series.....	137
7.2.5	Univariate and multivariate time series .....	138
7.2.6	Time series forecasting.....	138
7.2.7	Correlations .....	138
7.2.8	Interpolation.....	139
7.3	Timeseries forecasting approaches .....	139
7.3.1	Autoregressive model (AR).....	139
7.3.2	Moving Average model (MA) .....	140
7.3.3	Autoregressive Moving Average model (ARMA) .....	141
7.3.4	Auto Regressive Integrated Moving Average (ARIMA) .....	141
7.3.5	Exponential Smoothing .....	142
7.3.6	Multilayer Perceptron (MLP) .....	144
7.3.7	Long Short-term Memory (LSTM) .....	145
7.4	A methodology for time series forecasting of SaaS service factors .....	148
7.4.1	Overview of sentiment intensity forecasting methodology .....	148
7.4.2	Data preparation.....	149
7.4.3	Model selection.....	152
7.4.4	Parameter estimation .....	162
7.4.5	Model evaluation .....	174
7.5	Experiments and Evaluation results.....	175
7.5.1	Data preparation and Pre-processing.....	175
7.5.2	Model evaluation and comparative analysis.....	176
7.6	Conclusion .....	184

**Chapter 8: A Fuzzy-based Trust model for SaaS.....185**



8.1	Introduction: .....	185
8.2	Overview of the Fuzzy-based trust model for SaaS: .....	186
8.3	MDR weight metric .....	187
8.3.1	Maturity: 189	
8.3.2	Density: 189	
8.3.3	Reviewers:190	
8.4	Fuzzy-based Trust model:.....	192
8.4.1	SaaS service factors as Trust factors.....	194
8.4.2	Fuzzification of input variables and input membership functions.....	195
8.4.3	Fuzzy inference rules for the trust model .....	198
8.4.4	Rule aggregation and evaluation .....	201
8.4.5	Defuzzification to crisp trust values .....	202
8.5	Implementation and Case study .....	202
8.5.1	MDR scores calculation.....	203
8.5.2	Fuzzy inference.....	205
8.5.3	Experiments and Simulation results: .....	208
8.6	Conclusion .....	212
<b>Chapter 9: Recapitulation and future work.....</b>		<b>214</b>
9.1	Introduction .....	214
9.2	Recapitulation.....	215
9.3	Contribution of the thesis.....	216
9.3.1	Contribution 1: A novel framework that establishes the root of trust in SaaS using Architectural best practices.....	216
9.3.2	Contribution 2: An ensemble model that automatically identifies SaaS service factors	216
9.3.3	Contribution 3: An imputation method for missing sentiment scores of SaaS service factors	217
9.3.4	Contribution 4: A forecast model for SaaS service factors.....	218
9.3.5	Contribution 5: A dynamic trust model for SaaS .....	218
9.4	Future Work.....	219
9.5	Conclusion .....	220
<b>References.....</b>		<b>221</b>
<b>Appendices.....</b>		<b>230</b>

# List of Figures

Figure 1.1 Cloud computing conceptual reference model (Liu et al., 2011).....	2
Figure 3.1 Research approach.....	43
Figure 4.1. The proposed trust management framework for SaaS .....	48
Figure 4.2. Overview of classification model.....	50
Figure 4.3. Overview of the imputation model.....	52
Figure 4.4. Overview of the forecast model .....	54
Figure 4.5. Overview of the trust model .....	56
Figure 5.1 Sequence of steps for the SaaS service factors identification approach....	60
Figure 5.2. Number of samples per class before and after using SMOTE .....	69
Figure 5.3. Comparison of the training time for the classification models with their best parameters.....	87
Figure 5.4. Performance of the Ensemble model on training and test data with optimal parameters using SMOTE.....	90
Figure 6.1 Methodology for imputation of the missing service factors of SaaS .....	94
Figure 6.2 Sequence of steps in imputing the missing sentiment intensity values in a dataset .....	103
Figure 6.3 Steps in the validation process of Scheme A.....	107
Figure 6.4 Steps in the validation process of Scheme B.....	110
Figure 6.5 RMSE values averaged on all the datasets with different missing percentages of .....	114
Figure 6.6 RMSE scores comparison on all imputed datasets with different missing percentages .....	116
Figure 6.7 Adjusted R-squared scores comparison on all imputed datasets with different missing percentages .....	116
Figure 6.8 RMSE values averaged on all the datasets with different missing percentages of Scheme B.....	118
Figure 6.9 Cross-validation results from Linear and RBF kernels (RMSE) on T-NN imputed datasets.....	120
Figure 6.10 Cross-validation results from Linear and RBF kernels (Adjusted R-squared) on T-NN imputed datasets .....	121
Figure 6.11 RMSE scores comparison on all imputed datasets with different missing percentages during cross-validation .....	123
Figure 6.12 Adjusted R-squared scores comparison on all imputed datasets with different missing percentages during cross-validation .....	123

Figure 6.13 RMSE scores comparison on all imputed datasets with different missing percentages on test dataset.....	125
Figure 6.14 Adjusted R-squared scores comparison on all imputed datasets with different missing percentages on test dataset .....	126
Figure 6.15 RMSE scores comparison on imputed datasets using T-NN.....	127
Figure 6.16 Adjusted R-squared scores comparison on imputed datasets using T-NN.....	127
Figure 6.17 SVR and Baseline evaluation comparison on T-NN imputed and test datasets .....	128
Figure 6.18 SVR and Baseline evaluation comparison on T-NN imputed and test datasets .....	128
Figure 7.1 Components of time series .....	137
Figure 7.2 A multilayer perceptron with two hidden layers .....	145
Figure 7.3 Overview of a LSTM cell.....	146
Figure 7.4 Sequence of steps for the SaaS service factors forecasting methodology.....	148
Figure 7.5 Time series of each service factor .....	151
Figure 7.6 Time series transformation into supervised data .....	152
Figure 7.7 Stationary series after 1st-order differencing .....	154
Figure 7.8 Components of service factor E .....	155
Figure 7.9 Components of service factor L .....	155
Figure 7.10 Components of service factor S.....	156
Figure 7.11 Autocorrelations within the series for each service factor .....	160
Figure 7.12 Partial-autocorrelations within the series for each service factor.....	160
Figure 7.13 ARIMA model residuals.....	165
Figure 7.14 Autocorrelations in ARIMA model residuals.....	166
Figure 7.15 Partial autocorrelations in ARIMA model residuals .....	167
Figure 7.16 KDE of the ARIMA model residuals .....	168
Figure 7.17 Exponential smoothing model residuals.....	170
Figure 7.18 Autocorrelations in exponential smoothing model residuals .....	171
Figure 7.19 Partial autocorrelations in exponential smoothing model residuals.....	172
Figure 7.20 KDE of the exponential smoothing model residuals.....	172
Figure 7.21. MLP model summary.....	173
Figure 7.22 LSTM model summary.....	174
Figure 7.23 LSTM Single-step (Actual vs. Forecast).....	178
Figure 7.24 LSTM Multi-steps (Actual vs. Forecast).....	179
Figure 7.25 Multivariate models forecast errors.....	180
Figure 7.26 Univariate models forecast errors.....	180
Figure 7.27 Univariate vs. Multivariate models average forecast errors .....	181

Figure 7.28 Forecast errors comparison (Actual vs. Differenced).....	182
Figure 8.1 Overview of the framework for fuzzy-based trust model .....	187
Figure 8.2 Graphical representation of the individual metric in MDR.....	188
Figure 8.3 Flowchart for the sequence of tasks in fuzzy-based trust model.....	193
Figure 8.4 Main component of a Fuzzy Inference System .....	194
Figure 8.5 The input and output of Mamdani FIS .....	195
Figure 8.6 Input membership functions for all the input variables.....	197
Figure 8.7 Output membership functions for the fuzzy-based trust model .....	198
Figure 8.8 MDR scores for all the service factor from the SaaS dataset .....	205
Figure 8.9 Output membership functions for a given SaaS product's trust score and trust level .....	208
Figure 8.10 simulation results on the SaaS dataset and the percentage distribution of SaaS products based on their computed trust levels .....	209
Figure 8.11 Example output membership functions with trust levels and trust scores	210
Figure 8.12 Example output membership functions with trust levels and trust scores	210
Figure 8.13 Sentiment intensity scores for each service factor of a given SaaS product, used as trust factors for inputs in fuzzy trust model .....	211
Figure 8.14 The output trust scores at each time spot for which the input trust factors are observed.....	212

# List of Tables

Table 2.1 Trust modelling approaches .....	14
Table 2.2 Trust in different application domain .....	15
Table 2.3 Fuzzy trust models in cloud computing .....	32
Table 2.4 Critical analysis of Trust models in cloud computing .....	32
Table 5.1. Notations of the variables used in the experiments .....	61
Table 5.2. Service factors with associated tags.....	63
Table 5.3. Pre-processed review samples with associated labels .....	63
Table 5.4. Sparse matrix representation of review samples with TF-IDF scores for features .....	66
Table 5.5 Selected text classification approaches.....	69
Table 5.6. Selected parameters for tuning .....	70
Table 5.7. Dataset description with the number of reviews per service factor .....	79
Table 5.8. Optimal parameters selected after training .....	80
Table 5.9. Comparative analysis of the multiclass text classification approaches using Stratified k-folds and SMOTE .....	81
Table 5.10. Results from Nemenyi's post hoc test for classifier pairs .....	88
Table 5.11. Extended classification report of individual classes using ensemble model	88
Table 5.12. Confusion matrix for the ensemble classification model.....	89
Table 6.1. SaaS service quality pillars (service factors) .....	96
Table 6.2. Correlation values among the features.....	101
Table 6.3 Imputation approaches .....	105
Table 6.4. Datasets Description (scheme A).....	111
Table 6.5. RMSE values by increasing the threshold values (scheme A) .....	113
Table 6.6. RMSE scores with different missing percentages (scheme A).....	115
Table 6.7. Adjusted R-squared values with different missing percentages (scheme A)	115
Table 6.8. Datasets Description (scheme B).....	117
Table 6.9. RMSE values by increasing the threshold values (scheme B).....	119
Table 6.10. Cross-validation results with RMSE values using SVR on imputed datasets (scheme B) .....	122
Table 6.11. Cross-validation results with Adjusted R-squared values using SVR on imputed datasets (scheme B) .....	122
Table 6.12 Test results with RMSE values using SVR on imputed datasets (scheme B)	124

Table 6.13 Test results with Adjusted R-squared using SVR on imputed datasets (scheme B)	124
Table 6.14. SVR vs Baseline performance with RMSE and Adjusted R-Squared...	126
Table 6.15. Sample dataset with the sentiment intensity value SaaS service quality pillars (service factors)	130
Table 6.16. Correlation values among the features.....	130
Table 6.17. Service factors with correlation values and weights.....	130
Table 6.18. Forming a subset $D^-$ from the donor instances .....	131
Table 7.1. Augmented Dickey-Fuller Test (Significance level 0.05).....	153
Table 7.2. Granger Causality test (Significance level = 0.05).....	157
Table 7.3. Correlation results between service factors .....	158
Table 7.4. ARIMA optimal parameter for each service factor .....	163
Table 7.5. Exponential Smoothing model fit parameters .....	169
Table 7.6. MLP and LSTM parameters .....	173
Table 7.7. Parameters of the fitted MLP model.....	173
Table 7.8. Parameters of the fitted LSTM model .....	174
Table 7.9. SaaS sentiment intensity time series data description .....	176
Table 7.10. Single-step forecast errors .....	176
Table 7.11. Multi-steps forecast errors .....	177
Table 7.12. Average forecast errors.....	178
Table 8.1. SaaS service factors used as Trust factors for the fuzzy-based trust model	186
Table 8.2. Specifications of the triangular input fuzzy sets with their linguistic terms	196
Table 8.3. Specifications of the triangular output fuzzy sets with their linguistic terms	198
Table 8.4. Fuzzy inference rules with MDR scores for each Trust factor .....	199
Table 8.5. Samples of SaaS products with the sentiment intensity scores for each input Trust factor .....	203
Table 8.6. Details of individual metric used in MDR score calculation.....	204
Table 8.7. Fuzzy membership in each of the input membership function for the input variables.....	205
Table 8.8. Rule strengths and output fuzzy set association .....	206

# Glossary of terms

In this section, I present the key definition of the terms used in this chapter and rest of the thesis.

## **Cloud computing**

Cloud computing is a computing paradigm in which the computer resources are virtualized and are accessed over the internet as web services.

## **Cloud service**

A virtualized computing resource that is accessed over the internet.

## **Cloud service provider**

An entity that owns and provides cloud services.

## **Cloud service customer**

An entity that uses cloud services. The terms cloud service customers, cloud service users and cloud service consumers are used interchangeably in this thesis.

## **Cloud services delivery model**

The way a cloud service is delivered to the cloud customer i.e., as a virtualized infrastructure (Infrastructure as a Service), as a virtualized application development and deployment platform (Platform as a Service) and as a virtualized software (software as a Service).

## **Software as a Service**

A shared virtualized software that is used by the cloud service customers over the internet.

## **Cloud service selection**

The process of choosing cloud services from the available services.

## **Well-architected framework**

Frameworks proposed by Amazon web services (AWS, 2018) and Microsoft Azure (Azure, 2018) which are based on architectural best practices for cloud application design and operations.

## **Service factors**

The key factors for the optimal design and operations of cloud applications and services. The terms service aspect, service factor, trust factor, software quality pillars and pillars of well-architected framework are used interchangeably in this thesis.

**Customer review**

Textual feedback by a cloud service customer based on previous experience.

**Label**

Representation of a textual data into a given category.

**Data pre-processing**

The process of manipulating a textual data into an acceptable format as input for the next stage.

**Ensemble model**

A technical combination of two or more machine learning models.

**Sentiment intensity**

The strength of cloud service customer's opinion.

**Trust**

The belief in an entity.

**Trust model**

A formal representation of the belief in an entity. In this thesis, it is the belief in a Software as a Service.

**Trustworthiness level**

The linguistic representation of the level of belief in a Software as a Service.

**Trust value**

A numerical representation of the level of belief in a Software as a Service.

**Missing data**

An unobserved data during the data collection stage.

**Imputation**

The process of inserting a new value for an unobserved value of a given variable.



**Inference**

To deduce a missing observation by learning from the existing evidence following a reasoning technique.

**Threshold**

A numerical acceptance level for a given measurement.

**Time spot**

The instance of time for which an observation or forecast is referred to.

**Time slot**

An interval of time for which a set of observations or forecasts are divided.

**Time stamp**

A date or time label associated to a time spot.

**Forecasting**

The process of predicting a value for a given time spot in the future.

# Abstract

In recent years, cloud computing has gained much attention in research. Compared to its predecessor computing paradigms, cloud computing is more complex in its infrastructure and the nature of services it provides to its customers. Customers are reluctant to adopt cloud computing because of the ongoing issues such as data security, privacy and highly abstract nature of the cloud services. There is a need to facilitate consumer decisions when it comes to cloud services selection. Trust can play an important role in bringing the consumers closer to cloud computing as it has been successfully implemented in several domains. The complex nature of cloud services specially, the Software as a Service (SaaS), brings new challenges for trust modelling compared to the trust models in other domains. In the existing literature, there are a number of approaches proposed to model the trust in cloud services. However, the majority of existing research is focused on establishing the concept of trust and to assess the willingness of cloud service providers. Trust is multi-faceted, dynamic and is based on multiple factors. Besides cloud service providers willingness, the existing literature does not consider the SaaS architectures and their dependencies in their trust models. Besides that, computing trust in SaaS requires the information regarding all the trust factors. There is no approach in the existing literature to address the issue of missing information in the trust factors. Furthermore, the existing literature has not focused on forecasting the trust factors for the future time spots which is important in boosting customers confidence in the adoption of SaaS. The computation of trust in SaaS must adopt to time and should be reliable. The existing trust models for cloud services do not provide comprehensive solution that captures both dynamicity and reliability of trust factors.

Based on the shortcoming mentioned above, in this thesis I propose and develop a comprehensive trust management framework for SaaS. The proposed framework identifies the key trust factors from SaaS architectural best practices. Each component of the framework is designed to address the issues faced by trust models for SaaS. In the first component of the framework, an ensemble machine learning approach automatically categorizes and extracts the trust factors from SaaS customer reviews. An intelligent Threshold-based Nearest Neighbour (T-NN) method is proposed to impute the missing customer sentiments in SaaS service factors. The third component of the framework is designed to forecast the values of the SaaS service

factors. The forecast model is designed through a comprehensive study of the traditional and latest forecasting approaches. Finally, the last component of the framework, models the trust using an intelligent and dynamically weighted fuzzy trust model. Each component of the proposed trust management framework is validated against several well-known approaches during experiments.

This thesis will help and facilitate cloud services consumers in their decisions before utilizing cloud services and also the SaaS providers to improve different aspects of their offered services.