Neural Network with Multiple Training Methods for Web Service Quality of Service Parameter Prediction

Lov Kumar[1] Ashish Sureka[2]

[1]NIT Rourkela, India
Email: lovkumar505@gmail.com, skrath@nitrkl.ac.in

[2]ABB India, India
Email: ashish.sureka@in.abb.com

IC3 2017

Table of Contents

Table of Contents

Service Oriented Computing and Web Services

Selection of Best Web Service ?

Service-oriented computing : Reduction of heterogeneity between information systems through usage of standards [3][6][9]

Web services are web application components within the service-oriented and distributed computing paradigm

Web services providing similar functionality can have different QoS (Quality of Service) properties

Clients select the service for their requirements from amongst the similar functionality web services based on QoS attributes [10][14].

QoS Properties of Web Services

QoS for web services based on Source Code Metrics ?

QoS : Response time, throughput, availability, reliability, modularity and interoperability

QoS parameter values are sometimes not available and may change as the web service implementation and environment changes

Estimating the QoS parameters of web services through the source code implementing the web services can be useful

Provide guidance in terms of identifying regions of code that can improve the quality of web services.

Research Motivation

Source Code Metrics - QoS Parameters - Machine Learning

Investigate correlation between certain source code metrics and the QoS parameters

Examine whether the QoS parameters can be predicted using the source code metrics as features in a machine learning model

Correlation between 8 structural quality metrics, 19 data quality metrics and 21 procedural quality metrics with 12 QoS parameters

Investigate the application of neural networks with multiple training algorithms

Literature Survey

Kumar et al. [10]

Kumar et al. apply extreme learning machines with different types of kernel methods for estimating the quality of service parameters for web services [10].

Coscia et al. [11][5]

Coscia et al. demonstrate that several classic software engineering metrics can be used to predict important web service WSDL document quality attributes [11][5].

Literature Survey (Continued...)

Huang et al. [8]

Huang et al. present a method for web service selection based on using QoS properties as decisive factors from candidate services having similar functionality [8].

Cail et al. [4]

Cail et al. present an approach based on neural networks for modeling software reliability [4].

Novel and Unique Contributions

Correlation - Source Code Metrics and QoS

First study on investigating the correlation between 8 structural quality, 19 data quality, 21 procedural quality metrics of source code derived from web service interfaces with QoS

Neural network based model

We examine several variants of neural network algorithm by applying six different types of training methods

Empirical Analysis

Performance evaluation: Using MMRE, RMSE, PRED(25), comparison with previous work on similar problem.

Multi-Step Process

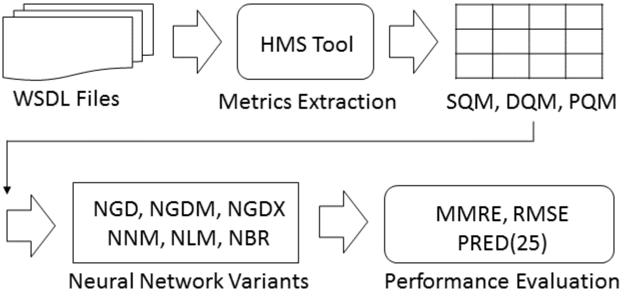

Figure: Source Code Metrics Computation, Neural Network Based Predictive Model Creation, Performance Evaluation and Benchmarking

Solution Approach and Framework

Multi-step Process

We use the Harry M. Sneed [12] tool to compute the structural, data and procedural quantity metrics from the WSDL interfaces.

Neural network classification for training a predictive model.

Neural network variants used in our experiments: NGD (gradient descent), NGDM (gradient descent with momentum), NGDX (variable learning rate gradient descent), NLM (levenberg-marquardt), NBR (bayesian regularization) and NNM (BFGS quasi-newton method)

We compare our proposed approach with a previous approach.

QWS Dataset

Experiments on freely and openly available dataset

We conduct experiments on QWS Dataset[a] which is publicly available and used by several web service researchers [1][2].

Kumar et al. [10] provide a list of 200 web services[b] used in their experiments and we use the same 200 web services in our experiments so that we can compare our results with the results obtained by Kumar et al. [10].

The minimum number of Java files are 5 and the maximum is 256.

[a] http://www.uoguelph.ca/~qmahmoud/qws/
[b] https://goo.gl/yrNyWz

Data Quality Metrics

Number of Data Structures	4.00	681.00	56.66	40.00
Number of Defined Definitions	1.00	479.00	19.18	1.00
Number of Data Variables declared	20.00	1762.00	109.43	70.00
Number of Data Variables inherited	0.00	0.00	0.00	0.00
Number of Data Constants/Enums declared	1.00	512.00	26.42	11.50
Number of Redefinitions (Unions)	0.00	0.00	0.00	0.00
Number of Arrays (Vectors)	0.00	31.00	1.27	0.00
Number of external Data Elements	0.00	0.00	0.00	0.00
Number of different Data Types used	4.00	983.00	87.06	67.00
Number of Data References	13.00	818.00	127.12	119.00
Number of Arguments / Input Variables	2.00	386.00	25.39	10.00
Number of Results / Output Variables	2.00	291.00	17.53	10.00

Figure: Refer to paper for complete details about all metrics

Structural Quality Metrics

	Min	Max	Mean	Median
	STRUCTURAL QUANTITY METRICS			
Number of Modules	2.00	2.00	2.00	2.00
Number of Includes	0.00	6.00	1.67	0.00
Number of Classes declared	2.00	487.00	35.22	29.00
Number of Classes inherited	0.00	0.00	0.00	0.00
Number of Operations declared	3.00	485.00	27.28	15.00
Number of Operations specified	3.00	485.00	27.28	15.00
Number of Procedures declared	0.00	0.00	0.00	0.00
Number of Interfaces declared	3.00	217.00	18.81	12.00

Figure: Refer to paper for complete details about all metrics

Descriptive Statistics Based Comparison

MMRE							
	Min	Max	Mean	Median	Std Dev	Q1	Q3
SQM	0.10	1.12	0.45	0.43	0.24	0.22	0.63
DQM	0.16	1.31	0.47	0.46	0.24	0.30	0.64
PQM	0.16	1.22	0.45	0.44	0.22	0.30	0.58
RMSE							
	Min	Max	Mean	Median	Std Dev	Q1	Q3
SQM	0.08	0.31	0.18	0.18	0.05	0.16	0.21
DQM	0.08	0.33	0.20	0.20	0.06	0.16	0.23
PQM	0.08	0.32	0.20	0.20	0.05	0.16	0.22

Descriptive Statistics Based Comparison

PRED(25)							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
SQM	0.17	0.94	0.55	0.49	0.23	0.38	0.70
DQM	0.17	0.92	0.51	0.45	0.23	0.33	0.66
PQM	0.18	0.92	0.52	0.50	0.22	0.37	0.64

Descriptive Statistics Based Comparison

Three Domain Metrics

Table (previous slide) displays the descriptive statistics of the performance of the three domain metrics for 12 QoS parameters and 6 neural network training methods.

The descriptive statistics is calculated from $72 = 12X6$ values. The performance comparison is done using MMRE, RMSE and PRED (25).

Descriptive Statistics Based Comparison (Continued..)

Three Domain Metrics

SQM (shaded in grey color) outperforms DQM and PQM in-terms of the three performance parameters.

Mean value of MMRE and RMSE for SQM is the lowest. The mean value of PRED(25) for SQM is highest whereas the median value is higher than the median value for DQM but lower than the median value for PQM.

DQM is the lowest performing metric. The mean of PRED(25) value of DQM is lower than the mean of PRED(25) value of PQM.

Six Neural Network Training Techniques

Table: Six Neural Network Training Techniques with Three Domain Metrics and 12 QoS Parameters

MMRE							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
NGD	0.17	1.22	0.51	0.49	0.26	0.30	0.66
NGDM	0.17	1.31	0.51	0.51	0.26	0.32	0.64
NGDX	0.17	1.20	0.49	0.47	0.25	0.30	0.62
NNM	0.11	0.93	0.43	0.43	0.21	0.23	0.60
NLM	0.10	0.78	0.40	0.38	0.20	0.21	0.50
NBR	0.10	0.80	0.40	0.39	0.20	0.21	0.49

Six Neural Network - PRED(25)

PRED(25)							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
NGD	0.18	0.91	0.48	0.41	0.23	0.30	0.65
NGDM	0.18	0.92	0.48	0.41	0.23	0.33	0.64
NGDX	0.18	0.91	0.50	0.45	0.23	0.34	0.66
NNM	0.17	0.93	0.53	0.49	0.23	0.37	0.66
NLM	0.19	0.94	0.57	0.54	0.21	0.45	0.68
NBR	0.17	0.94	0.57	0.55	0.21	0.44	0.67

Six Neural Network - RMSE

RMSE							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
NGD	0.10	0.32	0.21	0.20	0.06	0.17	0.24
NGDM	0.10	0.33	0.21	0.20	0.06	0.18	0.24
NGDX	0.09	0.33	0.20	0.20	0.06	0.17	0.24
NNM	0.08	0.28	0.18	0.18	0.05	0.16	0.22
NLM	0.08	0.25	0.18	0.19	0.05	0.15	0.21
NBR	0.08	0.27	0.18	0.18	0.05	0.16	0.21

Insights from Application of Neural Networks

For NLM the mean MMRE value is 0.40 and the median value is 0.38 which is lowest in comparison to the other neural network variants.

The RMSE and PRED(25) values for the six neural network variants shows that NLM outperforms all other classifiers.

The NLM model dominates in-terms of the mean values of MMRE, RMSE and PRED(25) performance metrics. [10].

NBR performance is nearest to NLM

Wilcoxon Signed Rank Test Results

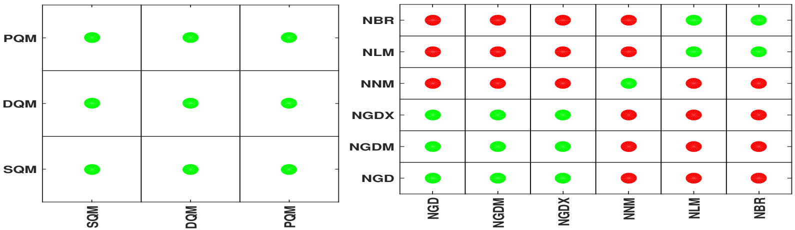

Figure: Wilcoxon Signed Rank Test Results (Statistical Significance Test). A red dot means that H_0 is rejected and a green dot means that H_0 is accepted.

Insights from Wilcoxon Signed Rank Test Results

Wilcoxon Signed-Rank Test is a pairwise comparison method and can be used to determine if the performance results is a significant indicator of one classifier being better than the other classifier [7]

There are a total of 3 pair-wise comparisons as we have 3 metrics. SQM is compared with DQM and PQM. DQM is compared with PQM.

Wilcoxon Signed-Rank Test generates a p-value based on comparing the performance of the two classifiers.

Insights from Wilcoxon Signed Rank Test Results

The null hypothesis H_0 is that there is no difference between the two classifiers being compared. The alternate hypothesis H_1 is that the metric characteristics is a significant influencer of the performance.

The adjusted p-value is computed by dividing 0.05 from ${}^{metrics}C_2 = 3$ and is a more stringent criteria.

There is a difference between the performance accuracy of the three metrics but the statistical test results indicates that the values obtained are not statistically significant and can be obtained by chance.

Wilcoxon Signed Rank - Neural Network Variants

In the case of neural network variants, the adjust p-value is even more stringent as the number of models to be compared are 6 in comparison to the number of domain metrics which is 3

Several red dots which shows that there is a statistically significant difference between the performances of the various neural network training methods

NLM model outperforms NNM, NGDX, NGDM, NBR and NGD in-terms of the MMRE, RMSE, PRED(25) values as well as from the perspective of Wilcoxon signed-rank test with a Bonferroni correction based comparison

Comparison with Benchmark

Table: Comparison of Approach with Kumar et al. ([10]) Method

MMRE							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
BMS	0.20	1.03	0.54	0.55	0.20	0.37	0.67
HMS	0.05	1.05	0.50	0.53	0.23	0.34	0.65
OOM	0.22	1.03	0.54	0.54	0.19	0.37	0.64
AM	0.05	1.18	0.51	0.48	0.24	0.35	0.69
PCA	0.06	0.96	0.51	0.50	0.21	0.34	0.68
RSA	0.30	1.12	0.58	0.55	0.21	0.37	0.68

Comparison with Benchmark

Table: Benchmarking and Comparison of Proposed Approach with Kumar et al. (refer [10]) Method

MMRE							
	Min	**Max**	**Mean**	**Median**	**Std Dev**	**Q1**	**Q3**
SQM	0.10	1.12	0.45	0.43	0.24	0.22	0.63
DQM	0.16	1.31	0.47	0.46	0.24	0.30	0.64
PQM	0.16	1.22	0.45	0.44	0.22	0.30	0.58

Comparison with Benchmark

Kumar et al. use a total of 37 source code metrics as predictors for the web services QoS parameters categorized into three classes: Chidamber and Kemerer Metrics (OOM), Harry M. Sneed Metrics (HMS) and Baski and Misra Metrics (BMS) [10].

They apply two different types of feature selection techniques: Principal Component Analysis (PCA) and Rough Set Analysis (RSA).

Table in previous slide reveals that the SQM metric outperforms all the 6 metrics used by Kumar et al. [10].

Comparison with Benchmark

The mean MMRE value of all the three metrics SQM, DQM and PQM outperforms all the six metrics used in previous work on the same dataset and same problem

The structural, procedural and data quantity metrics derived from the web service interfaces are better indicators than the classic object-oriented metrics

Table in previous slide reveals that the SQM metric outperforms all the 6 metrics used by Kumar et al. [10].

Final Conclusions and Takeaways

SQM metric outperforms DQM and PQM based on three performance metrics, 6 neural network training methods and 12 QoS parameters

NLM neural network model outperforms NNM, NGDX, NGDM, NBR and NGD in-terms of the MMRE, RMSE, PRED(25) values as well as from the perspective of Wilcoxon signed-rank test

Neural networks with structural, procedural and data quality metrics outperforms extreme learning based methods with classic object-oriented and Baski and Misra metrics.

References I

[1] Eyhab Al-Masri and Qusay H Mahmoud. Qos-based discovery and ranking of web services. In *Computer Communications and Networks, 2007. ICCCN 2007. Proceedings of 16th International Conference on*, pages 529–534. IEEE, 2007.

[2] Eyhab Al-Masri and Qusay H Mahmoud. Investigating web services on the world wide web. In *Proceedings of the 17th international conference on World Wide Web*, pages 795–804. ACM, 2008.

[3] Gustavo Alonso, Fabio Casati, Harumi Kuno, and Vijay Machiraju. Web services. In *Web Services*, pages 123–149. Springer, 2004.

[4] Kai-Yuan Cai, Lin Cai, Wei-Dong Wang, Zhou-Yi Yu, and David Zhang. On the neural network approach in software reliability modeling. *Journal of Systems and Software*, 58(1):47–62, 2001.

[5] José Luis Ordiales Coscia, Marco Crasso, Cristian Mateos, and Alejandro Zunino. Estimating web service interface complexity and quality through conventional object-oriented metrics. In *CIbSE*, pages 154–167, 2012.

References II

[6] Francisco Curbera, Matthew Duftler, Rania Khalaf, William Nagy, Nirmal Mukhi, and Sanjiva Weerawarana. Unraveling the web services web: an introduction to soap, wsdl, and uddi. *IEEE Internet computing*, 6(2):86–93, 2002.

[7] Salvador Garcia and Francisco Herrera. An extension on statistical comparisons of classifiers over multiple data sets"for all pairwise comparisons. *Journal of Machine Learning Research*, 9(Dec):2677–2694, 2008.

[8] Angus FM Huang, Ci-Wei Lan, and Stephen JH Yang. An optimal qos-based web service selection scheme. *Information Sciences*, 179(19):3309–3322, 2009.

[9] L. Kumar, S. Rath, and A. Sureka. Using source code metrics to predict change-prone web services: A case-study on ebay services. In *2017 IEEE Workshop on Machine Learning Techniques for Software Quality Evaluation (MaLTeSQuE)*, pages 1–7, 2017.

[10] Lov Kumar, Santanu Kumar Rath, and Ashish Sureka. Predicting quality of service (qos) parameters using extreme learning machines with various kernel methods. In *4th International Workshop on Quantitative Approaches to Software Quality*, page 11, 2016.

References III

[11] José Luis Ordiales Coscia, Marco Crasso, Cristian Mateos, and Alejandro Zunino. Web service interface quality through conventional object-oriented metrics. *CLEI Electronic Journal*, 16(1):5–5, 2013.

[12] Harry M Sneed. Measuring web service interfaces. In *Web Systems Evolution (WSE), 2010 12th IEEE International Symposium on*, pages 111–115. IEEE, 2010.

[13] Mie Mie Thet Thwin and Tong-Seng Quah. Application of neural networks for software quality prediction using object-oriented metrics. *Journal of systems and software*, 76(2):147–156, 2005.

[14] Liangzhao Zeng, Boualem Benatallah, Anne HH Ngu, Marlon Dumas, Jayant Kalagnanam, and Henry Chang. Qos-aware middleware for web services composition. *IEEE Transactions on software engineering*, 30(5):311–327, 2004.