

Applying data mining in telecommunications

Radosavljevik, D.

Citation

Radosavljevik, D. (2017, December 11). *Applying data mining in telecommunications*. Retrieved from https://hdl.handle.net/1887/57982

Version: Not Applicable (or Unknown)

License: License agreement concerning inclusion of doctoral thesis in the

Institutional Repository of the University of Leiden

Downloaded from: https://hdl.handle.net/1887/57982

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle http://hdl.handle.net/1887/57982 holds various files of this Leiden University dissertation.

Author: Radosavljevik, D.

Title: Applying data mining in telecommunications

Issue Date: 2017-12-11

Curriculum Vitae

Dejan Radosavljevik was born in 1975 in Skopje, Macedonia. After graduating from a BSc program in Computer Science at the University of Ss. Cyril and Methodius in Skopje in 2001, he has worked as a software developer for several Macedonian companies. In 2009 he completed a Master's degree in ICT in Business with cum laude distinction at Leiden University with a thesis on Prepaid Churn Modeling Using Customer Experience Management Key Performance Indicators. Since then he has worked in multiple positions related to artificial intelligence, data mining and data science at T-Mobile Netherlands B.V., in parallel to working on this PhD research at Leiden University. He currently holds the position of Lead Data Scientist within T-Mobile Netherlands.

List of Figures

1.1	CRISP-DM Process Model for Data Mining	9
2.1 2.2	Customer Experience Framework for Mobile Telecommunications Coefficient of Concordance	22 24
2.3	Coefficient of Concordance of predictors grouped in group 1 for experiment A	30
2.4		32
3.1	Telecom call graph	41
3.2	Initial energy of the simple and extended propagation technique	43
3.3	Spreading activation in a weighted graph	44
3.4	Call Graph Details	46
3.5	Implementation scenarios	48
3.6	Gain and Lift chart of all models	50
4.1	Gain Charts of Models Used	60
5.1	Actual Load vs. Linear approximation	73
5.2	Communication Graph of the Tools used	
6.1	Overview of the Service Revenue Forecasting Process	
6.2	The ETL Process in KNIME using RJDBC	92
6.3	Modeling Workflow in KNIME	96

132 LIST OF FIGURES

List of Tables

1.1	Mapping of the Focus of the Thesis Chapters to the Stages of the CRISP-DM process	11
2.1 2.2	Sample size, churn rate and CoCs in experiments A, B1a, B1b and C Grouping of variables of Model A_Incl_CEM $\ldots \ldots \ldots \ldots$	29 31
3.1 3.2	Social network features used in the extended tabular churn models Coefficient of Concordance of the scoring and propagation models	42 51
4.1 4.2 4.3 4.4	List of contractual, demographic and CDR based features List of network quality features per category	57 58 59 61
5.1	List of Input Parameters	71
5.2	Regression Modeling Results for Downlink Load (DL) for Country Operator 1	78
5.3	Regression Modeling Results for Uplink Load (UL) for Country Operator 1	78
5.4	Regression Modeling Results for Downlink Load (DL) for Country Operator 2	78
5.5	Regression Modeling Results for Uplink Load (UL) for Country Operator 2	78
5.6	Regression Modeling Results for Downlink Load (DL) for Country Operator 3	79
5.7	Regression Modeling Results for Uplink Load (UL) for Country Oper-	79 79
5.8	ator 3	
5.9	Operator 4	79 - 2
	ator 4	79

134	LIST OF TABLES
101	Elet et mille

6.1	Inputs used for creating service revenue models	94
6.2	Algorithm performance	100
6.3	Modeling Service Revenue Components	101