Classification of Body Regions based on MRI Log Files

Nadine Kuhnert[1,2], Oliver Lindenmayr[1,2] and Andreas Maier[1]

[1] Pattern Recognition Lab, FAU Erlangen Nuremberg, Germany
[2] Siemens Healthcare GmbH, Erlangen, Germany

Abstract. Every Siemens Magnetic Resonance Imaging (MRI) system consistently writes events into log files while the system is running. The log files and their contents are constantly refined by software developers. This results in different information contents depending on the software version. One information that is missing in some log files is the examined body region. As the body region is crucial for usage analysis, we used pattern recognition methods to estimate the examined body region for software versions not logging it automatically. We learned the examined body region from a set of used MRI acquisition parameters such as grid and voxel size and could classify body region information with a classification rate up to 94.7%. We compared Bayesian Network augmented Naïve Bayes, Decision Trees, and Neural Networks, and found Neural Networks resulting in the best classification rate.

Keywords: Classification, Data Mining, Pattern Recognition, Log File Analysis, System Usage

1 Introduction

Medical imaging devices not only produce medical images but also generate log files. The log file content is developed with every released software version. Thus, the log files provide insight into the detailed, subsequent events of a running system. This information can constitute tremendous added value e.g. for predictive maintenance, trouble shooting, and usage analysis. With the customer's approval, the log files are systematically analysed also for Magnetic Resonance Imaging (MRI) systems. MRI systems produce images of the anatomy and physiological processes of the body. Mostly, only specific body parts are of interest and thus, only those body parts are examined. Especially for usage analysis and understanding requirements to the hardware and software of the system, knowing the examined body region is essential. However, this information is logged automatically only by systems with a software version released later than 2010. Thus, many currently used MRI scanners do not deliver the examined body part automatically. We want to solve this problem using Machine Learning (ML), learn from scans with given examined body parts, and classify those examinations where the body region is unknown.

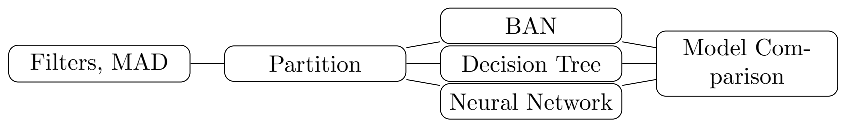

Fig. 1: Process chain illustrating the filters, transformation, and partition applied to the training data. The preprocessed data presents the input for the methods BAN, Decision Tree Learning, and Neural Networks. In the final step their results are tested on the validation set and compared.

Data mining is omnipresent in web usage and more and more data mining is also applied in medical applications. For instance, Asaro et al. [1] used medical record access logs for the identification of non-characteristic user behaviour in order to detect unauthorized access to medical records. Moreover, Gallagher et al. [2] use log file information of medical systems to address security issues as well as operational aspects. Thus, they implemented an audit system that tracks the access information per user. Data mining, as well as ML, describes generating additional knowledge of the given information [3]. Many ML algorithms are shared with concepts used in pattern recognition (PR) and are presented by Duda et al. [4]. Both, ML and PR, include the approach of Neural Networks which process, dispose, and generate new information from given data, similar to the human brain [5]. Furthermore, Decision Tree Learning is a predictive modelling approach by partitioning the example space. E.g. Patil et al. [6] use Decision Trees, Support Vector Machines, and k-Nearest Neighbour for the classification of examination types in interventional X-ray imaging. Also Chen et al. [7] use path analysis, sequential pattern discovery and other methods to clinical information system log files in order to discover patient-specific clinician information needs. With this gained knowledge, they displayed the most decisive patient information and thus, could improve the user interface as well as workflow efficiency. Due to its simplicity and good results Naive Bayes is used in many discussions, compared with other approaches, and modified in different ways [8][9].

2 Materials and Methods

In this section, we present the ML approaches Augmented Naïve Bayesian Network (BAN), Decision Tree Learning, and Neural Networks and their solutions to our problem implemented in SAS® Enterprise Miner™. Figure 1 illustrates the main steps.

We used 70% of the data set with available body region for training and 30% for validation. Our data set contains 16147 scans recorded in log files from 105 MRI scanners belonging to six different system types. The scans have been executed during a time frame of three days on scanners all over the world.

Variable	Distribution (%)
Head	39.53
Spine	15.85
Heart	12.80
Abdomen	8.96
Knee	8.42
Brain	4.96
Shoulder	3.97
Extremity	2.78
Pelvis	2.72

Table 1: Class-wise distribution of sample data.

2.1 Preprocessing

All three methods were applied to the same training and validation set after preprocessing the data. First of all, we neglected empty examples as well as outliers using the Median Average Deviation [10] and only selected scans with known body region. This yields 15036 scans. Subsequently, we transformed the data using quantile binning and applied stratified partitioning such that we use 70% (10527) for training and 30% (4509) for validation. Afterwards, three in the following described classification methods have been applied to the preprocessed data. We present in Table 1 the distribution of body regions in percent in our training set which constitute the classes. Furthermore, Table 2 shows the parameters given in the log file data which we used for the classification.

2.2 Augmented Naïve Bayesian Network

The first approach is based on the state-of-the-art classifier naïve Bayes and Bayesian networks. Introduced by Friedman et al. [11] as augmented naïve Bayesian network (BAN), the classifier combines the simplicity of naïve Bayes and the ability to cope with independence. A Bayesian network consists of random variables, represented by nodes, and edges connecting the nodes. The edges stand for conditional dependency between two nodes, whereas a naïve Bayesian network only connects every input variable to the attribute class due to the assumption of independence. On the contrary, BAN may have edges connecting the variables with non-class attributes in addition [12].

In this work the maximum number of edges one node is connected with others is restricted to five. This allows a fair amount of dependence between variables but also keeps the search space low. The network structure is built on independence tests (constraint-based) and such that the structure fits the training data best (score-based) [13]. According to the result shown in Table 3, the used classifier has a misclassification rate of 7.1%.

Interval Features	Min	Max	Mean	Skewness
Echo Time	0	107.0	24.967	1.429
Inversion Time	80	1918.0	439.779	1.359
Repetition Time	2.5	5960.0	1196.475	1.316
Flip Angle	0	180.0	88.379	0.187
Number of Images	1	30.0	6.853	1.641
Slice Measurement Duration	15	512047.0	87005.770	0.916
Total Scan Time	1	714.0	121.133	1.015
Slice Thickness	0.5	13.6	5.198	0.599
Nominal Features	**Levels**			
Field Of View	33			
Coil	33			
Contrast Bolus	2			
Acquisition Matrix	33			
Patient Position	4			

Table 2: MRI parameters used as features for classification of body regions.

2.3 Decision Tree Learning

Another method with which we used to learn the missing body part information from is called Decision Tree Learning. It learns and makes decisions based on a decision tree. A decision tree consists of one root node describing the input and several branches and leaves. Branches constitute decision rules and leaves represent classes containing a subset of the input. In our implementation, decisions for the division are based on variance reduction. Each node is divided into two branches successively. The maximum depth is ten and the minimal number of entries in one node is set to five [14]. The resulting confusion matrix is given in Table 4 and shows a misclassification rate of 7.0%.

2.4 Neural Networks

We classified the MRI scans to the nine classes of body regions also using a Neural Network. During the training phase the Neural Network learns and adapts the weights in between the units such that the correct class is assigned depending on the scan parameters [15]. We chose three layers, maximal iteration of 300, and used 14139 weights. We show the class-wise confusion matrix in Table 5 indicating a classification rate of 94.7%. Furthermore, it lists the number of scans assigned to the respective body region and the actual one. It illustrates that the differentiation of brain and head poses the hardest decision with 96 and 42 wrong assignments, and the predominant main diagonal presenting the correct classification.

For training the Neural Network, we used the MRI parameters and settings for the individual scans. The chosen coil constitutes the main indicator with an

Actual Class	Predicted Class: Spine	Shoulder	Pelvis	Knee	Heart	Head	Extremity	Brain	Abdomen
Spine	645	7	5	0	0	4	7	16	6
Shoulder	10	120	0	1	0	0	0	1	0
Pelvis	5	2	66	0	0	0	1	1	9
Knee	0	0	3	271	0	0	8	0	0
Heart	0	0	1	0	401	1	0	0	4
Head	22	32	7	10	6	1799	22	230	3
Extremity	1	0	1	0	0	0	56	0	3
Brain	14	7	1	0	0	13	0	369	4
Abdomen	11	5	8	6	14	7	4	1	258

Table 3: Result of the augmented naïve Bayesian network in form of a class-wise confusion matrix.

importance of 66.6%. Table 6 lists all used features and their importance to the weights of the Neural Network.

Moreover, we performed a ten-fold cross validation for training and validation, separately. We used a data set containing 754707 examples. Figure 2 illustrates the resulting misclassification rates using a bar plot. The horizontal axis carries the 10 folds named F1 to F10, whereas, the vertical axis depicts the misclassification rate in percent. The best rate was reached by F7 with 5.63%, the worst by F5 with 5.88% misclassified scans. The number of examples used for training and validation for every fold is given in Table 7. In addition to the misclassification rate (MCR), it shows the root mean square error (RMSE) with a maximal value of 0.085.

3 Discussion

We classified 16147 MRI scans by their examined body regions using BAN, Decision Tree Learning, and Neural Network. The Neural Network delivered the best classification rate of the validation set of 94.7%. BAN classified the scans resulting in a classification rate of 92.9%, we reached 93.0% with the approach of Decision Tree Learning. We showed in Table 6 that the chosen coil for the examination is a main contributor to the correct classification. However, the best classification requires the incorporation of all nine parameters. In future work, the Neural Network should be tested and developed further for other incomplete data sets and parameters in order to counteract inconsistent log file data sets due to different software versions of the systems.

	Predicted Class								
Actual Class	**Spine**	**Shoulder**	**Pelvis**	**Knee**	**Heart**	**Head**	**Extremity**	**Brain**	**Abdomen**
Spine	657	1	3	0	0	2	1	10	16
Shoulder	9	116	0	1	1	1	0	1	3
Pelvis	7	1	63	0	2	0	0	0	11
Knee	0	0	0	280	0	2	0	0	0
Heart	4	0	1	0	400	0	0	0	2
Head	10	0	0	4	1	2074	0	36	6
Extremity	3	0	2	10	0	1	46	0	5
Brain	10	0	1	0	3	104	0	283	7
Abdomen	13	4	7	0	10	1	1	1	271

Table 4: Result of the Decision Tree Learning in form of a class-wise confusion matrix.

	Predicted Class								
Actual Class	**Spine**	**Shoulder**	**Pelvis**	**Knee**	**Heart**	**Head**	**Extremity**	**Brain**	**Abdomen**
Spine	674	0	1	0	0	3	4	5	3
Shoulder	7	119	0	0	1	1	0	2	2
Pelvis	1	1	75	0	0	0	1	0	6
Knee	0	0	0	280	0	0	2	0	0
Heart	0	0	3	0	402	0	0	1	1
Head	13	0	0	3	0	2072	0	42	1
Extremity	5	0	2	5	0	0	52	0	3
Brain	8	1	0	0	0	96	0	302	1
Abdomen	7	0	1	0	5	0	0	2	293

Table 5: Result of the Neural Network in form of a class-wise confusion matrix.

Variable	**Importance (%)**
Coil	66.6
Flip Angle	26.2
Number of Images	25.9
Patient Position	19.1
Repetition Time	16.6
Echo Time	14.4
Slice Thickness	13.5
Slice Measurement Duration	12.9
Field Of View	12.7
Acquisition Matrix	11.9
Total Scan Time	8.9
Inversion Time	4.6
Contrast Bolus	4.0

Table 6: The feature importance in percent to the weights of the Neural Network.

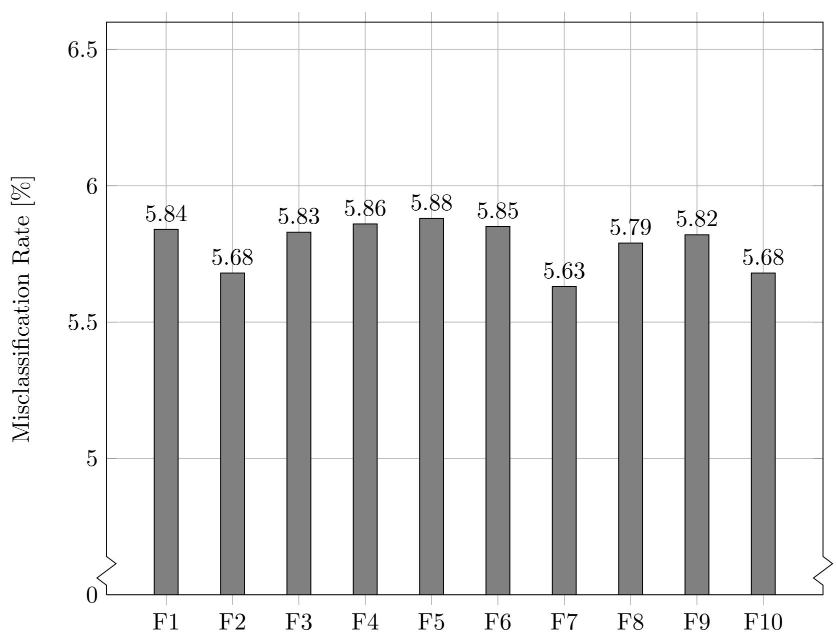

Fig. 2: Misclassification rates of a 10-fold cross-validation method.

Folds	**Training**	**Validation**	**MCR**	**RMSE**
Total	528691	226016	5.89	0.086
1	476033	203328	5.84	0.085
2	475697	203491	5.68	0.084
3	475927	203387	5.83	0.085
4	475845	203369	5.86	0.085
5	475967	203330	5.88	0.085
6	476038	203408	5.85	0.085
7	475714	203409	5.63	0.083
8	475810	203345	5.79	0.085
9	475677	203620	5.82	0.085
10	475511	203457	5.68	0.084

Table 7: Data sets of 10-fold cross-validation. Number of used examples for training and validation are given next to the misclassification rate (MCR) and the root mean square error (RMSE) per run.

References

1. Asaro, P. and Ries, J.: Data Mining in Medical Record Access Logs. Proceedings of the AMIA Symposium, Washington, DC, USA, p. 855 (2001)
2. Gallagher, R., Sengupta, S., Hripcsak, G., Barrows, R., Clayton, P.: An Audit Server for Monitoring Usage of Clinical Information Systems (1997)
3. Witten, I., Frank, E.: Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann (2005)
4. Hart, P., Stork, D., Duda, R.: Pattern Classification. John Willey & Sons (2001)
5. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall PTR, 2nd Edition (1998)
6. Patil, M., Patil, R., Krishnamoorthy, P., John, J.: A Machine Learning Framework for Auto Classification of Imaging System Exams in Hospital Setting for Utilization Optimization. 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA 2423–2426 (2016)
7. Chen, E., Cimino, J.: Automated discovery of patient-specific clinician information needs using clinical information system log files. AMIA Annual Symposium Proceedings, Washington, DC, USA 145–9 (2003)
8. Domingos, P., Pazzani, M.: On the optimality of the simple Bayesian classifier under zero-one loss. Machine learning 29, 103–130 (1997)
9. McCallum, A., Nigam, K.: A comparison of event models for naive bayes text classification. AAAI-98 workshop on learning for text categorization vol. 752, 41–48 (1998)
10. Hampel, F.: The influence curve and its role in robust estimation. Journal of the American Statistical Association vol. 69, 383–393 (1974)
11. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian Network Classifiers. Machine Learning vol. 29, 131–163 (1997)
12. Keogh, M., Pazzani, M.: Learning Augmented Bayesian Classifiers: A Comparison of Distribution-based and Classification-based Approaches (1999)
13. SAS Institute Inc: SAS®Enterprise Miner™14.1 Reference Help (2015)
14. Ville, B. and Neville, P.: Decision Trees for Analytics Using SAS Enterprise Miner (2013)
15. Han, J. and Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, 2nd Edition (2006)