

PAPER • OPEN ACCESS

## Using Machine Learning to predict survival in patients with brain metastases after Gamma Knife radiosurgery

To cite this article: G Vazhenin *et al* 2019 *J. Phys.: Conf. Ser.* **1205** 012058

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

# Using Machine Learning to predict survival in patients with brain metastases after Gamma Knife radiosurgery

G Vazhenin<sup>1</sup>, S Banov<sup>3</sup>, A Golanov<sup>2</sup>, P Ryabov<sup>1</sup>, V Kostjuchenko<sup>3</sup> and A Dalechina<sup>3</sup>

<sup>1</sup>National Research Nuclear University "MEPhI", Kashirskoe highway, 31, Moscow, Russia

<sup>2</sup>N.N. Burdenko Neurosurgical Institute, 1-st Tverskoy-Yamskoy lane, 13/5, Moscow, Russia

<sup>3</sup>Moscow Gamma Knife Center, 1-st Tverskoy-Yamskoy lane, 13/5, Moscow, Russia

avdalechina@gmail.com

**Abstract.** In this work machine learning approach was used to predict the patients overall survival after Gamma Knife radiosurgery. We constructed the regression and multiclass classification models to predict the time interval from the onset of the oncological disease to the unfavourable outcome and the patient's survival class. The models were built on data of 916 patients with 26 different primary features. The train set included patients with known clinical outcomes (445 patients). The median deviation in determining of the time from the onset of cancer to the date of death was 1.4 months. According to the regression model the most significant feature was the largest volume of the lesion by the date of the first radiosurgery. The most important feature was identified as the time interval from the date of birth to the onset of the oncological disease. The mean accuracy, according to the confusion matrix was 0,76. The accuracy of the study can be improved by the increasing of the patients number and completeness of the patient data.

## 1. Introduction

The quantity of detected cases of oncology diseases is being on arise all over the world. According to the World Health Organization's data the oncology diseases are within the leading mortality factors. On a conservative estimate the brain metastases are in progress in 8 - 10 % of cancer patients, what determines principal clinical symptomatology. The key method to treat the multiple brain metastases combines the various types of radiotherapy with modern pharmacotherapy (immunotherapy, target therapy etc.).

At the moment the most efficient and secure method in case of limited brain metastases is considered to be radiosurgery. The essential feature of radiosurgery is the locality of radiation exposure which is substantially decreasing the toxicity. It enables to undertake the radiosurgery procedure multiply controlling the growth of the lesions which could be re-appeared in 50 % of the cases.

It is extremely important for cancer research to predict outcomes after a given therapy and detect the key features from complex datasets and identify the prognostic factors effecting patients survival. But traditionally prediction has been performed by clinicians using only statistically significant clinical



factors such as age, performance status, diagnosis and etc. Every patient is complex and presents with a great number of clinical factors that may or may not be relevant to survival [1].

Machine Learning (ML) was proposed as an alternative method of data analysis. It is a powerful tool to overcome the limitations of classical statistical methods while analyzing and interpreting massive amounts of data [2]. ML enables computers to learn from previous data to make accurate predictions on new data.

The purpose of this study is to apply the ML to predict survival in patients with brain metastases (BS) after Gamma Knife radiosurgery (GKRS). We applied ML algorithms to the tasks of classification and regression to predict overall survival. ML methods were implemented using open-source Python libraries: scikit learn, catboost, pandas, numpy, scipy, pylab.

## 2. Materials and methods

### 2.1. Patients selection

We retrospectively evaluated the total data of 916 patients with brain metastases undergone the stereotactic Gamma Knife radiosurgery treatment in Moscow Gamma Knife Center (Burdenko neurosurgical institute) within the period from 2005 to 2017. These patients had 26 different primary features: Karnofsky Performance Scale (KPS), oncological diagnosis, primary tumor histology, number of brain tumors and etc. The data could be divided into 4 classes: real, categorical, date, text feature (Table 1).

**Table 1.** Primary features<sup>a</sup>.

Categorical variables	Real variables	Date	Text
1. Sex	10. Number of GKRS	15. Date of birth	26. Histology
2. OD	11. KPS on the date of the first RS	16. Diagnosis date	
3. Activating mutations	12. Number of brain tumors on the date of the first RS	17. Date of the primary tumor surgery	
4. EM by the date of the first RS	13. Total volume of the BM on the date of the first RS	18. Date of the BM appearance	
5. Pharmacotherapy by the date of the first RS	14. Maximum volume of the brain tumor by the date of the first RS	19. WBRT date	
6. IP		20. Date of the BM neurosurgery	
7. LR treatment after the first RS		21. Date of the first GKRS	
8. LR after the first RS		22. Date of the LR appearance after the first RS	
9. DM after the first RS		23. Date of the DM appearance after the first RS	
		24. Date of the last follow-up	
		25. Date of the patient's death	

<sup>a</sup> Abbreviations used in the table: OD - oncological diagnosis, EM - extracranial metastases, RS - radiosurgery, IP - intracranial progression, LR - local recurrence, DM - distant metastases, WBRT - whole brain radiation therapy.

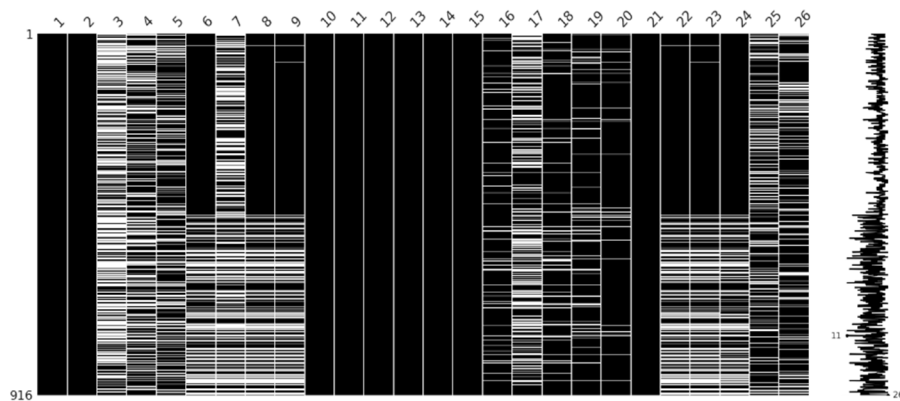
The data were divided into training and test sets. The target variable was time measured in months from the oncological diagnosis date (the onset of the oncological disease) to the date of the patient's

death. There was information about the date of death of 479 patients in the data. But the train set contained only 445 patients with information about both the date of death and the oncological diagnosis date. The test set in turn did not include any information about the date of death for these patients or they had been alive by the date of the last observation. Predictions were made for the test set.

For the classification task we divided the training set into 3 classes: 0 (297 patients) – time interval from the first RS to the death less than 1 year, 1 (95 patients) – time interval from the first RS to the date of death from one to two years, 2 (87 patients) – time interval from the first RS to the date of death more than 2 years or the patient was alive by the date of the last observation. Thereby, in classification, alive patients by the date of the radiosurgery were included into analysis.

## 2.2. Data pre-processing

Missing data is a common problem in real ML tasks. Our dataset had 24 % of the missing values. The visualization of the missing data is presented in the Fig.1.



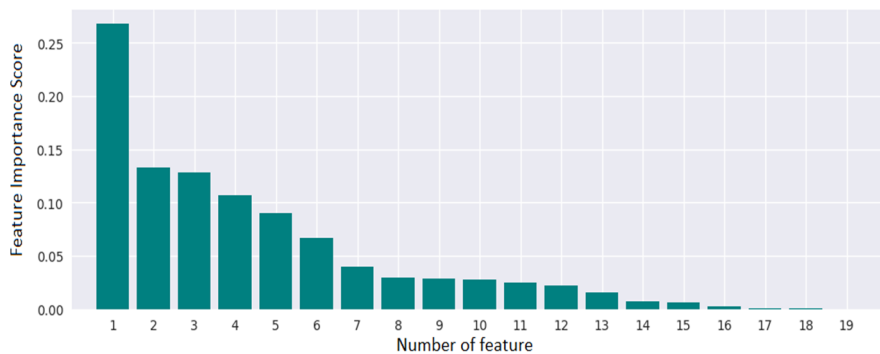
**Figure1.** Missing data visualization. White fields indicate missing values.

The missing values were replaced with 0 (in case of the categorical features) and with median value (in case of the missing values in the date features). Categorical features were transformed to the binary format. The “date” features were transformed into the differences between dates. Quantities of words, letters, numbers were calculated for the text feature “Histology”. The appearance of the most frequently used words of the “Histology” feature was normalized to binary features. The data filled less than 10 % were excluded from the analysis.

## 2.3. Data pre-processing

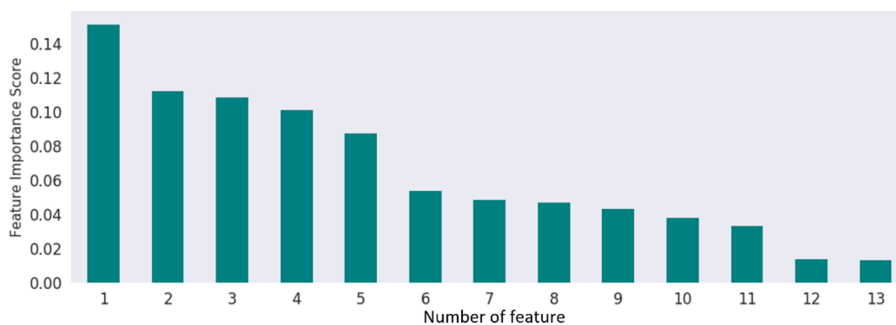
Feature selection is the process of identifying and removing from the training data set irrelevant and redundant features. Training time increases with number of features. Besides model has risk of overfitting with increasing number of features. Using Random Forest classifier and greedy algorithm the most relevant features were selected [3].

For the regression task, we identified 19 features: 1- maximum volume of the brain tumors on the first RS date, 2 - total volume of the BM by the date of the first RS , 3 - time interval from the birth to the onset of the oncological disease, 4 - absence of the intracranial progression, 5 - KPS by the date of the first RS, 6 - IR treatment after the first RS, 7 -GKRS number, 8 - IR without treatment after the first RS, 9 - LR, 10 - absence of the EM by the date of the first RS, 11- DM, 12 - chemotherapy, 13 - male, 14 - EM by the date of the first RS, 15 - Breast cancer, 16 - Melanoma, 17 - NSCLC (non-small cell lung cancer), 18 - IR treatment – WBRT, 19 - Kidney cancer. The bar chart is presented in the Fig.2.



**Figure 2.** Feature Importance bar chart (regression model).

For the classification task the relevant features were: 1 - time period from the birth to the onset of the oncological disease, 2 - total volume of the BM on the date of the first RS, 3 - maximum volume of the brain tumors on the first RS date, 4 - time interval from the disease onset to the BM appearance, 5 - time from the BM appearance to the first GKRS, 6 - IR treatment - RS, 7 - number of BM treated with RS, 8 - KPS by the date of the first RS, 9 - number of words in the “histology feature”, 10 - GKRS number, 11 - quantity of numbers in the “histology feature”, 12 - Breast cancer, 13 - Melanoma (Fig.3).



**Figure 3.** Feature Importance bar chart (classification task).

#### 2.4. Tuning model hyperparameters and training

Predictive models were trained using Gradient Boosting Regressor and Gradient Boosting Classifier in cases of regression and classification tasks respectively [4]. Using variational principle, we found the optimal Gradient Boosting hyperparameters for the classification task: minimum number of samples required to split an internal node = 200, minimum number of samples required to be at a leaf node = 30, maximum depth of the tree = 7, number of the features to consider when looking for the best split = 7. For the regression task default scikit-learn hyperparameters were used.

### 3. Results

Predictive models were constructed to predict overall survival in patients with BM using ML approach.

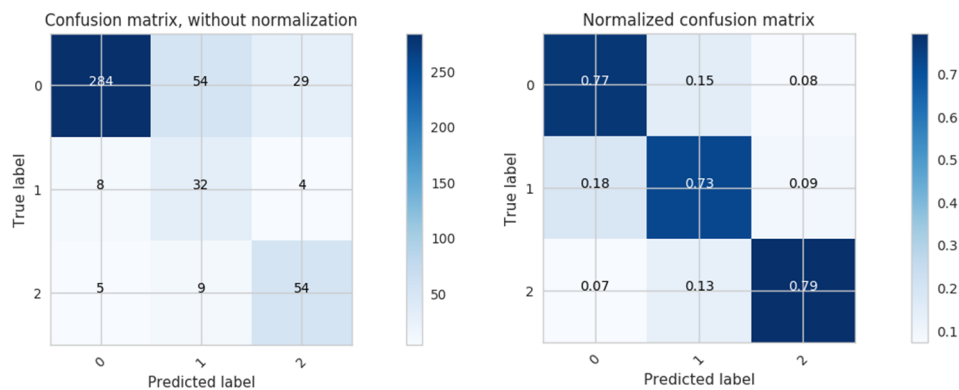
It was founded that the most significant features to predict the time interval from the oncological diagnosis date to the date of the patient’s death were the largest volume of the lesion by the moment of the first RS, total volume of the BM by the date of the first RS, time interval from the birth to the onset of the oncological disease, absence of the intracranial progression, KPS by the date of the first RS.

Five most significant features in classification were: time period from the date of birth to the onset of the oncological disease, total volume of the BM by the date of the first RS, maximum volume of the

brain tumors by the first RS date, time interval from the disease onset to the BM appearance, time from the BM appearance to the first GKRS.

The 5-fold cross validation technique was used to estimate the accuracy of the ML models. The mean and the median errors of the regression predictive model were 1.8 and 1.4 months, respectively.

The mean accuracy of the classifier was 0.72. For the classification problem the confusion matrix was used to evaluate the quality of the output of the classifier (Fig.4.). It allowed to evaluate correctness of the model for each class (mean accuracy - 0,76).



**Figure 4.** Confusion matrix (non-normalized and normalized)

The mean accuracy according to the confusion matrix was roughly equal to the mean accuracy of the classifier without reference to the classes. It demonstrated that we succeeded in correcting of the imbalanced classes and the imbalance of the data had no significant influence on the prediction accuracy. The presented confusion matrix has showed the approximate equality of the accuracies ref to each separate class, that is also an indicator of the well-balanced classes.

#### 4. Conclusion

This research is just a first step to apply the ML algorithms to prediction of treatment results in case of patients with BS in the Moscow Gamma Knife Center. It has been found that the most important features are the largest volume of the lesion by the moment of the first radiosurgery (the regression task) and the time interval from the date of birth to the onset of the oncological disease (the classification task). The prediction error of the regression model was 1,4 months and the mean accuracy of the classification was 0,76. The accuracy of the study can be improved by the increasing number of patients and completeness of the patient data.

#### References

- [1] Kondziolka D *et al* 2014 *J. Neurosurg.* **120** 24–30
- [2] Celticki E 2018 *Turk Neurosurg.* **28** 167 - 73
- [3] Kam Ho T 1995 ICDAR '95 Proceedings of the Third International Conference on Document Analysis and Recognition **1** 278
- [4] Friedman J 2001 *Stat.* 1189–32.