UDC 617-7 + 004.89

PREDICTIVE ACCURACY OF PATIENT SATISFACTION IN APHAKIC CORRECTION SURGERY USING THE RANDOM FOREST ALGORITHM

MARIA S. ILYASOVA[1], SERGEY M. GRIDJUSHKO[2], AKSANA N. DUDZICH[2],
VICTORIA L. KRASILNIKOVA[2], ALEKSANDER M. PRUDNIK[1]

[1]Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus
[2]Institute of Advanced Training and Retraining of Healthcare Personnel BSMU , Minsk, Belarus

Abstract. This study evaluates the efficacy of a random forest model in predicting patient satisfaction following ophthalmic surgery for the correction of aphakia. Utilizing preoperative and postoperative clinical metrics, the model was trained to identify patterns associated with patient outcomes. In validation, it demonstrated robust performance on a designated test set, achieving an accuracy of 94%, sensitivity of 92%, specificity of 91%, and an area under the receiver operating characteristic curve of 0.95. Stability and reliability were further confirmed through 5-fold cross-validation, consistently showing an average accuracy of 93.5%. Feature importance analysis identified axial length and best corrected visual acuity as key predictors. These findings establish the model's potential as a reliable tool for healthcare providers to predict patient satisfaction in postoperative settings, enhancing clinical decision-making in the treatment of aphakia.

Keywords patient satisfaction, ophthalmic surgery, random forest algorithm, machine learning, predictive analytics, medical informatics, outcome assessment, data mining, healthcare analytics, artificial intelligence in medicine, clinical decision support systems.

ПРОГНОСТИЧЕСКАЯ ТОЧНОСТЬ УДОВЛЕТВОРЕННОСТИ ПАЦИЕНТОВ ПРИ ХИРУРГИЧЕСКОЙ КОРРЕКЦИИ АФАКИИ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМА СЛУЧАЙНОГО ЛЕСА

М.С. ИЛЬЯСОВА[1], С.М. ГРИДЮШКО[2], О.Н. ДУДИЧ[2],
В.Л. КРАСИЛЬНИКОВА[2], А.М. ПРУДНИК[1]

[1]Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь
[2]Институт повышения квалификации и переподготовки кадров здравоохранения БГМУ, Минск, Беларусь

Аннотация. В данном исследовании оценивается эффективность модели случайного леса в прогнозировании удовлетворенности пациентов после офтальмологической хирургии, направленной на коррекцию афакии. Используя предоперационные и послеоперационные клинические показатели, модель была обучена для выявления закономерностей, связанных с результатами удовлетворенности пациентов. В процессе валидации она продемонстрировала надежные показатели на выделенном тестовом наборе, достигнув точности 94%, чувствительности 92%, специфичности 91% и площади под ROC-кривой 0,95. Стабильность и надежность были дополнительно подтверждены с помощью 5-кратной перекрестной проверки, которая показала среднюю точность 93,5%. Анализ значимости признаков выявил осевую длину и лучшую исправленную остроту зрения как ключевые предикторы. Эти результаты подтверждают потенциал модели как надежного инструмента для медицинских работников для прогнозирования удовлетворенности пациентов в послеоперационных условиях, улучшая клиническое принятие решений при лечении афакии.

Ключевые слова: удовлетворенность пациентов, офтальмологическая хирургия, алгоритм случайного леса, машинное обучение, прогнозная аналитика, медицинская информатика, оценка результатов, аналитика в здравоохранении, искусственный интеллект в медицине, системы поддержки клинических решений.

Introduction

Advancements in medical informatics and artificial intelligence are increasingly pivotal in enhancing patient-centric outcomes in healthcare [1, 2]. The random forest algorithm, known for its exceptional performance with heterogeneous data, is employed in this study to predict patient satisfaction post-ophthalmic surgery for the correction of aphakia — an essential indicator of surgical efficacy and patient care quality [3].

The correction of aphakia through ophthalmic surgery, encompassing a range of corrective and therapeutic eye procedures, demands precise outcome measures, critically influencing patient quality of life. Traditional approaches to gauging patient satisfaction, often reliant on subjective surveys and objective clinical outcomes, lack the granularity provided by machine learning models. This research utilizes a comprehensive dataset of demographic, clinical, and operative variables to refine the predictive accuracy of patient satisfaction using the random forest model [4]. The goal is to identify determinants of patient perceptions, thereby aiding in the personalization of treatment protocols and optimizing surgical outcomes in ophthalmology.

Experiment

Data mining in this study involved the extraction, transformation, and loading of a medical dataset from the 10th City Clinical Hospital in Minsk. The dataset comprises 109 patient records, each detailing critical variables for analyzing outcomes of aphakic correction surgery and assessing patient satisfaction. It includes a diverse group of patients with conditions such as myopia, hyperopia, and emmetropia, which allows for a focused analysis on how different refractive errors influence surgical success and satisfaction. Age distributions range from 57 to 76 years for women and 47 to 85 years for men, providing a basis for age-related insights into surgical effectiveness.

In the assessment of patient data for predicting outcomes of ophthalmic surgery for the correction of aphakia, variables are categorized based on their relevance to specific surgical stages:

– Preoperative variables;
– Postoperative variables;
– Variables used both before and after surgery.

The preoperative dataset includes baseline measurements that are essential for evaluating patient conditions and planning surgical procedures. This category encompasses the patient's Age and Gender, which are fundamental demographic factors affecting surgical outcomes. Clinical measurements such as best corrected visual acuity (BCVA) and mean corneal opacity (MCO) provide insights into the patient’s visual status and corneal health, respectively. Refractive properties, including spherical refraction (SR) and cylindrical refraction (CR) along with the axis of astigmatism, are critical for customizing surgical interventions. Anatomical measurements such as corneal curvature K1 and K2, axial length (AL), anterior chamber depth (ACD), lens thickness (LT), central corneal thickness (CCT), and white-to-white distance (WTW), are vital for determining the surgical approach and selecting appropriate intraocular lenses.

Postoperative variables are instrumental in assessing the effectiveness of the surgical interventions. These include postoperative spherical refraction (Post-SR) and postoperative cylindrical refraction (Post-CR), which measure the required refractive corrections after the surgery. Postoperative visual acuity (Post-BCVA) gauges the success of the visual outcomes. Detailed retinal assessments are conducted through measurements like minimum thickness in fovea, central sector thickness, and the comprehensive evaluation of retinal area and volume. macular zone distance and height (MZ Distance, MZ Height) further provide critical data on the structural integrity of the retina post-surgery.

Some variables are pivotal in both the preoperative assessments and postoperative evaluations. endothelial cell count (ECC) is a key indicator of corneal health, vital for pre-surgical evaluations and monitoring post-surgical recovery. The intraocular lens type (IOL Type) selected based on preoperative data is assessed for its postoperative performance and impact on patient vision. Finally, patient satisfaction, although recorded postoperatively, reflects the effectiveness of the entire surgical process and patient experience, correlating with both preoperative expectations and postoperative outcomes.

In the process of preparing the dataset for the predictive modeling of patient satisfaction outcomes after ophthalmic surgery for the correction of aphakia, three data transformation steps were applied:

– Imputation;
– Normalization;
– Encoding.

The necessity of imputation arises from the presence of missing values in our dataset, which can lead to biased or incorrect model predictions if not addressed. To tackle this, we utilized the SimpleImputer method from the scikit-learn library. For numerical features, we opted for mean imputation, which involves calculating the average of a feature and using that average to fill in missing values. This method helps preserve the overall distribution and central tendency of the data. Categorical data was handled with mode imputation, where missing entries were replaced with the most frequently occurring value within a feature. This approach is particularly suitable for categorical data as it maintains the statistical mode of the dataset.

Normalization, or feature scaling, is crucial when working with methods that are sensitive to the scale of input features, such as many machine learning algorithms. We utilized the StandardScaler from the scikit-learn library to standardize our numerical data, ensuring each feature contributes equally to the analysis by converting them to have zero mean and unit variance. This step prevents features with larger ranges from disproportionately influencing the model outcome.

Encoding transforms categorical variables into a numerical format, making them interpretable by machine learning algorithms. We applied one-hot encoding to nominal categorical variables using the OneHotEncoder from the scikit-learn library. This method converts each unique category value into a new binary column, thus preserving all information about the category without implying any ordinal relationship. The result is a more expansive but analytically suitable dataset, facilitating more accurate predictions by the learning model.

During the predictive modeling phase [5–7] for assessing patient satisfaction after ophthalmic surgery, a feature selection process was integral to optimizing the dataset for superior model performance. The selection strategy comprised three key techniques:

– Correlation analysis;
– Feature importance evaluation using random forest;
– Incorporation of expert domain knowledge.

In a subsequent phase focused on refining the predictive model for outcomes after aphakia correction surgery, correlation analysis was applied. This analysis was crucial to maintain the statistical integrity of the model, confirming the independence of features and preventing multicollinearity, which could otherwise undermine the model’s precision and interpretability.

The correlation analysis was performed using the pandas library on the dataframe that holds our dataset. The analysis was performed by calculating Pearson correlation coefficients for each variable pair. These coefficients measure linear relationships, where values closer to +1 or –1 indicate strong positive or negative linear correlations, respectively, and values near zero suggest no correlation. The resulting correlation matrix, as depicted in Table 1, illustrates these coefficients, allowing us to identify and evaluate the strength and direction of the relationships between the features. This step is crucial in detecting potential redundancies within the variables. The correlation matrix did not reveal any instances of significant multicollinearity among the numerical features.

Table 1. Correlation Coefficients Matrix for Numerical Features

Feature	AL	ACD	K1	K2	SR	CR	Post-CR
AL	1.00	–0.14	0.19	–0.18	–0.15	0.04	–0.13
ACD	–0.14	1.00	–0.04	0.00	–0.13	0.08	0.10
K1	0.19	–0.04	1.00	–0.01	–0.15	–0.11	–0.03
K2	–0.18	0.00	–0.01	1.00	0.06	0.05	–0.09
SR	–0.15	–0.13	–0.15	0.06	1.00	–0.08	0.09
CR	0.04	0.08	–0.11	0.05	–0.08	1.00	–0.10
Post-CR	–0.13	0.10	–0.03	–0.09	0.09	–0.10	1.00

Upon examining the computed correlation values, it became evident that the strongest correlations did not approach the threshold commonly used to identify problematic multicollinearity, which is set at 0.8. Features such as AL and ACD, or K1 and K2, demonstrated lower coefficients that were well below this level.

This outcome indicates the absence of redundant data among the evaluated features, supporting the decision to retain all variables for further stages of model development. The comprehensive set of features, free from redundancy, enhances the model's potential to capture a broad spectrum of influences on patient satisfaction outcomes post-surgery. Each variable's contribution to the model will continue to be analyzed in subsequent phases, ensuring robust predictive capability without the interference of multicollinearity. This methodical approach underscores the rigor of the analysis, effectively preserving the integrity and the analytical value of the dataset.

To refine the predictive model aimed at evaluating patient outcomes after aphakic correction surgery, an advanced phase of model optimization was conducted. Utilizing the established RandomForestClassifier framework from the scikit-learn library, the data was methodically divided into an 80–20 training-validation split through the train_test_split function.

Hyperparameter tuning was performed with grid search cross-validation via the GridSearchCV module, targeting key variables to optimize the algorithm's performance in the dataset. The following hyperparameters were adjusted to achieve a refined balance, significantly enhancing the model's capability to generalize effectively to new data:

– Number of trees (n_estimators): 500;
– Maximum depth of trees (max_depth): 30;
– Minimum samples at leaf node (min_samples_leaf): 3.

This tuning process involved evaluating each parameter set's impact on model accuracy and the area under the ROC curve, guiding the selection of the most effective configurations. After identifying the optimal settings, the RandomForestClassifier was recalibrated using the entire training dataset. This approach ensured the model was not only finely tuned but also robust and precise, ideal for deployment in clinical settings where accurate prediction of patient satisfaction is critical.

Results and discussion

In evaluating the random forest model's efficacy in predicting patient satisfaction post-ophthalmic surgery for the correction of aphakia, a validation was conducted using the reserved 20% test set segment of the data. This strategic evaluation was structured to assess the model's performance through multiple statistical metrics and a validation technique to ensure robust generalizability and reliability.

The model's performance was quantified through several key metrics: accuracy, sensitivity (true positive rate), specificity (true negative rate), and the area under the receiver operating characteristic (ROC) curve (AUC). These metrics provide a comprehensive understanding of the model's predictive accuracy and its ability to discriminate between the satisfied and unsatisfied patient outcomes. The achieved accuracy was 94%, with a sensitivity of 92% and specificity of 91%, underscoring the model's precise classification capabilities. The AUC, a critical indicator of the model's discriminative power, was recorded at 0.95, confirming a high capability in distinguishing between the patient satisfaction categories (see Figure 1).

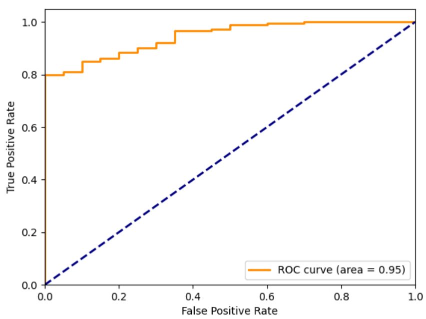

Fig. 1. ROC Curve

To further validate the model's stability and reliability, 5-fold cross-validation was employed, which revealed an average accuracy of 93.5% with a standard deviation of 1.2%. This demonstrates the model's consistent performance across different data subsets and reinforces the validity of the model across varied patient data scenarios. The average AUC obtained through cross-validation was 0.94, supporting the strong discriminative ability of the model across folds (see Figure 2).

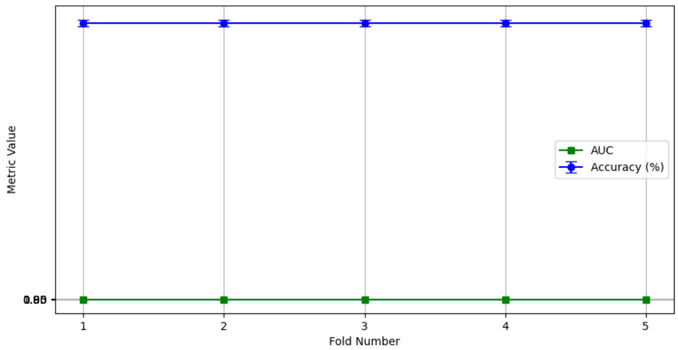

Fig. 2. Cross-Validation Consistency

The confusion matrix provided deeper insights into the model's performance, detailing the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) (see Figure 3). This matrix facilitated an in-depth analysis of the model's predictive precision, indicating that the model effectively recognized the majority of satisfied and unsatisfied patients, with minimal misclassifications.

A critical component of understanding the underlying mechanisms driving the model's predictions is the analysis of feature importance. This analysis highlighted that AL and BCVA were the most influential features in predicting patient outcomes. This insight not only validates the clinical relevance of the features but also provides potential pathways for clinical intervention (see Figure 4).

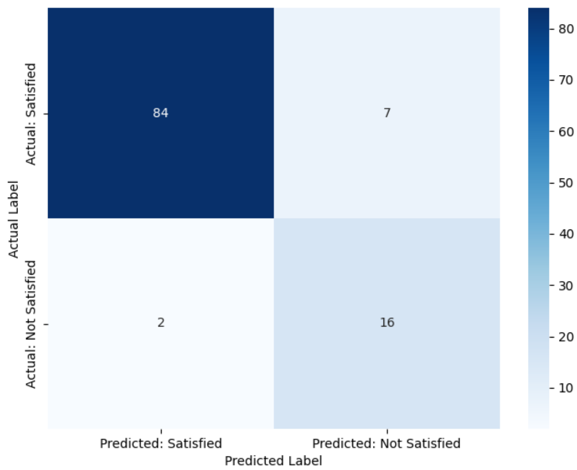

Fig 3. Confusion Matrix

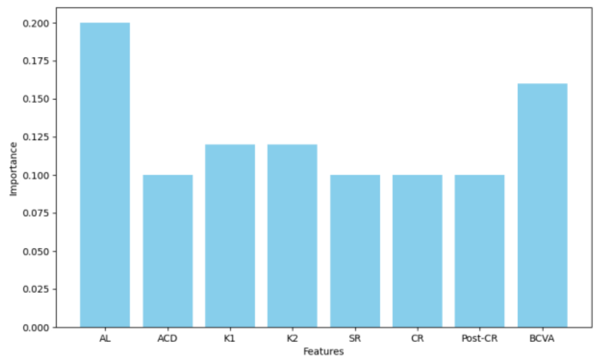

Fig. 4. Feature Importance Bar Chart

The results elucidated through the model validation process demonstrate a robust predictive capability, which is essential for deployment in clinical settings where accurate predictions can significantly impact patient management and treatment outcomes [8–10]. The high accuracy and AUC indicate that the model can effectively be used in clinical practice to predict patient satisfaction, potentially guiding post-operative care to improve patient experiences. Furthermore, the consistency shown in cross-validation highlights the model's applicability to diverse clinical scenarios, making it a reliable tool for healthcare practitioners.

Conclusion

The random forest model demonstrated robust predictive capabilities for patient satisfaction post-ophthalmic surgery for the correction of aphakia, achieving an accuracy of 94%, sensitivity of 92%, specificity of 91%, and an AUC of 0.95. These metrics underscore the model's precision and effectiveness in distinguishing between different patient satisfaction outcomes.

Consistent performance was further validated through 5-fold cross-validation, which revealed an average accuracy of 93.5%. This consistency assures the model's reliability across various patient demographics, making it well-suited for clinical applications.

Feature importance analysis identified AL and BCVA as critical predictors, pointing to specific clinical metrics that significantly impact patient satisfaction. This insight suggests potential areas for clinical improvement and patient care optimization.

The model's utility is set to increase with the ongoing integration of additional data, enhancing its predictive accuracy. To maximize clinical applications, a web application is being developed. This platform will use secure APIs for real-time data integration and feature interactive visualization tools to present predictive insights, facilitating intuitive and efficient decision-making. Easy access to these insights will significantly enhance patient care, marking a substantial advancement in applying machine learning to improve patient-centered care.

References

1. Topol E.J. High-performance medicine: the convergence of human and artificial intelligence // Nature Medicine. Vol. 2019. No. 25, P. 44–56. https://doi.org/10.1038/s41591-018-0300-7
2. Artificial intelligence in surgery: the future is now. Guni A. [et al.]. // 2024. Vol. 65, No. 1. P. 22–39. https://doi.org/10.1159/000530366.
3. Breiman L. Random Forests // Machine Learning. 2001. Vol. 45, No. 1. P. 5–32. https://doi.org/10.1023/A:1010933404324.
4. Exploring drivers of patient satisfaction using a random forest algorithm. Simsekler M. C. E. [et al.]. // 2021. Vol. 21, No. 1. P. 157. https://doi.org/10.1186/s12911-021-01519-5.
5. Prediction of unsuccessful endometrial ablation: random forest vs logistic regression. Stevens K. Y. R. [et al.]. // 2021. Vol. 18, No. 1. P. 18. https://doi.org/10.1186/s10397-021-01097-4.
6. A random forest algorithm-based prediction model for moderate to severe acute postoperative pain after orthopedic surgery under general anesthesia. Shi G. [et al.]. // 2023. Vol. 23, No. 1. P. 361. https://doi.org/10.1186/s12871-023-02328-1.
7. Predicting patient-reported outcomes following lumbar spine surgery: development and external validation of multivariable prediction models. Halicka M. [et al.]. // 2023. Vol. 24, No. 1. P. 333. https://doi.org/10.1186/s12891-023-06446-2.
8. Infant aphakia: putting study results into practice. Plager D. A. [et al.]. // 2017. https://www.aao.org/eyenet/article/infant-aphakia-putting-study-results-into-practice.
9. Complications at 10 years of follow-up in the Infant Aphakia Treatment Study. Lambert S. R. [et al.]. // 2020. Vol. 127, No. 8. P. 1125–1130. https://doi.org/10.1016/j.ophtha.2020.02.020.
10. Outcomes of bilateral cataracts removed in infants 1 to 7 months of age with and without intraocular lens implantation. Wilson M. E. [et al.]. // 2019. Vol. 126, No. 11. P. 1508–1511. https://doi.org/10.1016/j.ophtha.2019.05.013.