Improving clinical refractive results of cataract surgery by machine learning

Martin Sramka, Martin Slovak, Jana Tuckova, Pavel Stodulka

1 Department of Circuit Theory / Faculty of Electrical Engineering, Czech Technical University of Prague, Prague, Czech Republic
2 Research and development department, Gemini Eye Clinic, Zlin, Czech Republic
3 Department of Ophthalmology / Third Faculty of Medicine, Charles University Prague, Prague, Czech Republic

Corresponding Author: Martin Sramka
Email address: sramka@gemini.cz

Aims. To evaluate potential of Support Vector Machine Regression model (SVM-RM) and Multilayer Neural Network Ensemble model (MLNN-EM) as an intraocular lens (IOL) power calculation improvement for clinical workflow.

Background. Current IOL power calculation methods offer limited accuracy and especially in eyes with an unusual ocular dimension the accuracy may decrease. In case of an improperly calculated power of the IOL in cataract or refractive lens replacement surgery, there is a risk of re-operation or further refractive correction. This may potentially induce complications and discomfort to the patient.

Methods. A dataset containing information about 2194 eyes was obtained using data mining process from Electronic Health Record (EHR) system database of Gemini Eye Clinic. The dataset was optimized and split into Selection set (used in design of models and training), and Verification set (used in the evaluation). Set of mean prediction errors and distribution of predicted refractive error were evaluated for both models and Clinical results.

Results. Both models performed significantly better for the most evaluated parameters compared to the clinical results. There was no significant difference between both evaluated models. In the ± 0.50 D prediction error category both SVM-RM and MLNN-EM were slightly better than Barrett Universal II formula (which is often presented as the most accurate calculation formula).

Conclusion. In comparison to the method currently used in a clinical setting, both SVM-RM and MLNN-EM has achieved significantly better results in IOL calculations and therefore there is a strong potential to improve clinical cataract refractive outcomes.
Improving Clinical Refractive Results of Cataract Surgery by Machine Learning

Martin Sramka¹,², Martin Slovak², Jana Tuckova¹, Pavel Stodulka²,³

¹ Department of Circuit Theory / Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic.
² Research and development department, Gemini Eye Clinic, Zlin, Czech Republic.
³ Department of Ophthalmology / Third Faculty of Medicine, Charles University, Prague, Czech Republic.

Corresponding Author:
Martin Sramka
Email address: sramka@gemini.cz

Abstract

Aims. To evaluate potential of Support Vector Machine Regression model (SVM-RM) and Multilayer Neural Network Ensemble model (MLNN-EM) as an intraocular lens (IOL) power calculation improvement for clinical workflow.

Background. Current IOL power calculation methods offer limited accuracy and especially in eyes with an unusual ocular dimension the accuracy may decrease. In case of an improperly calculated power of the IOL in cataract or refractive lens replacement surgery, there is a risk of re-operation or further refractive correction. This may potentially induce complications and discomfort to the patient.

Methods. A dataset containing information about 2194 eyes was obtained using data mining process from Electronic Health Record (EHR) system database of Gemini Eye Clinic. The dataset was optimized and split into Selection set (used in design of models and training), and Verification set (used in the evaluation). Set of mean prediction errors and distribution of predicted refractive error were evaluated for both models and Clinical results.

Results. Both models performed significantly better for the most evaluated parameters compared to the clinical results. There was no significant difference between both evaluated models. In the
± 0.50 D prediction error category both SVM-RM and MLNN-EM were slightly better than Barrett Universal II formula (which is often presented as the most accurate calculation formula).

**Conclusion.** In comparison to the method currently used in a clinical setting, both SVM-RM and MLNN-EM has achieved significantly better results in IOL calculations and therefore there is a strong potential to improve clinical cataract refractive outcomes.

**Introduction**

Cataract surgery is the principal lens replacement refractive surgical procedure performed in adults and one of the most commonly performed surgical procedures today (Abell & Vote, 2014; Frampton et al., 2014; Wang et al., 2017b). Every year, more than 11 million eyes undergo IOL implantation worldwide. The World Health Organization has estimated that number of cataract blind people will increase from 10 million in 2010 to 40 million in 2025 as the population grows (Pascolini & Mariotti, 2012). Phacoemulsification with IOL implantation is currently the most common method of treating cataracts and many refractive vision errors for which other conventional methods are not suitable (Linebarger et al., 1999). The ultimate goal is to achieve complete postoperative independence of ocular correction. Since significant developments of cataract and refractive surgeries over the past twenty years we are now even closer in meeting this target, although there are still areas we can improve.

The quality of the patient's post-operative vision depends on a correct choice of the IOL optical power, which influences the residual post-operative refraction. Improvement of refractive result of the cataract surgery is a challenge for the IOL manufactures but also for the methods used in the calculation of suitable IOL lens power.

To achieve accurate IOL calculation a series of scientific and therapeutic approaches needs to be made. This includes an accurate examination of the reason for the vision loss (Yamaguchi, Negishi & Tsubota, 2011), preoperative ocular surface preparation, patient visual preferences, eye biometric measurements (Astbury & Ramamurthy, 2006; Shammas & Shammas, 2015), precise eye surgery and IOL positioning (Thulasi, Khandelwal & Bradley Randleman, 2016), and last but not least is to have an accurate IOL power calculation method (Norrby, 2008; Lee et al., 2015).

To determine the optimal IOL power the calculation formulas are used. These formulas use data from preoperative measurements, examinations and IOL parameters, which may all influence the overall outcome. The calculation formulas can be divided into Refraction, Regression, Vergence, Artificial intelligence and Ray tracing categories based on their calculation method (Koch et al., 2017).
Currently, the most commonly used formulas are from the Vergence formula category. Although their accuracy achieving only ± 0.5 diopter (D) from the intended target refraction in 60-80% of eyes (Melles, Holladay & Chang, 2018). Their accuracy decreases even further in the eyes with non-standard biometric features such as eyes with short or long axial lengths (Abulafia et al., 2015; Shrivastava et al., 2018).

The only currently used IOL calculation approach using Artificial Intelligence is the Hill-RBF formula, which has reported accuracy of 91% of the eyes within ± 0.5 D range from the intended target refraction (Haag-Streit AG Koeniz, 2017). However, there is a number of papers that indicate that Hill-RBF accuracy is not significantly different from the Vergence formula category (Kane et al., 2017; Roberts et al., 2018; Shajari et al., 2018). Unfortunately, there is no publication of the Hill-RBF principle in any peer-reviewed scientific journal, so the only information about the principle itself can be obtained from freely available resources on the Internet. Based on this information, it is possible to find out that the Hill-RBF core is Radial Basis Function and that the algorithm was trained on the data of more than 12,000 eyes. There is no evidence of what specific machine learning method is used (Hill; Snyder, E.; The American Society of Cataract and Refractive Surgery; Haag-Streit AG Koeniz, 2017). Radial Basis Functions are used in many applications in the field of biomedical engineering (Le & Ou, 2016a,b).

This paper aims to describe the methodology of selecting and optimizing dataset for SVM-RM and MLNN-EM training, to describe a methodology for evaluating the accuracy of the model, to evaluate SVM-RM and MLNN-EM for IOL power prediction and to compare accuracy of both models with the current calculation method used in the clinical practice.

Support vector machine (SVM) is a supervised machine learning method serving mainly for classification and, in our case, for regression analysis. The aim of this algorithm is to find a hyperplane that optimally splits the feature space so that training data belonging to different classes lie in the separable spaces (Smola & Schölkopf, 2004). To find such a hyperplane on non-linear data, a kernel trick is used to transform data from the original feature space into a higher dimension space where it is already linearly separable (Herbrich, 1999; Jap, Stöttinger & Bhasin, 2015). SVM regression introduces an epsilon-insensitive loss function that is taken into account when minimizing the error through hyperplane optimization. SVM find their application for example in the field of financial forecasting (Trafalis & Ince, 2000), travel time prediction (Chun-Hsin Wu et al., 2003), flood forecasting (Yu, Chen & Chang, 2006) and genetics (Le et al., 2019).

Multilayer Neural Networks (MLNN) are known for their exceptional ability to approximate continuous functions (Mongillo, 2011; Wu et al., 2012) and has been widely used in function approximation, prediction, and classification (Park & Sandberg, 1991; Girosi, 1992; Clarke & Burmeister, 1997; Ferrari & Stengel, 2005).
The MLNN network consists of a collection of inputs and processing units known as neurons which are organized in the network layers. Neuron parameters are set up by the training process described by (Kurban & Beşdok, 2009). The training process is determined by minimizing an error function that measures the degree of success in the recognition of a given number of training patterns (Lampariello & Sciandrone, 2001).

Materials & Methods

The project research can be structured into three main parts: Dataset Preparation, Model Design & Training and Evaluation (Fig. 1).

Data preparation focuses on the methods used in data collection, storage in EHR database, data mining, cleaning and optimization in order to obtain suitable dataset for training and evaluation. Incorrect integration of these processes could lead to a degradation of data sources and a distortion of the quality of results.

Model design and training part focuses on a set-up of suitable SVM-RM and MLNN-EM models and their training using the dataset.

Evaluation part describes the outcome measures and how the data were analyzed.

This study used data of the patients who underwent cataract or lens replacement surgery from December 2014 to November 2018, at Gemini Eye Clinic, Czech Republic. This study was approved by the Institutional Ethics Committee of the Gemini Eye Clinic (IRB approval number 2019-04) and adhered to the tenets of the Declaration of Helsinki. As this was anonymous retrospective data collection study, written patient consent was waived by IRB.

Data acquisition

Data were acquired, recorded and stored by skilled staff in the central EHR at Gemini Eye Clinic. Data were usually entered before the surgery and at follow up visits and post-operative examinations.

The preoperative patient evaluation included distance objective refraction ($R_{x_{pre}}$), distance subjective refraction, mean keratometry (K), anterior chamber depth (ACD), axial length of the eye (AL), uncorrected distance visual acuity (UDVA), corrected distance visual acuity (CDVA), slit lamp examination, retinal examination and contactless intra-ocular pressure examination. Anterior and posterior segment evaluations and biometry measurements were conducted on all patients in the dataset. All biometry examinations (K, ACD, AL) were conducted using Carl Zeiss IOL Master 500 (Carl Zeiss, Jena, Germany) (Chen, Hirnschall & Findl, 2011). All measurements of objective refraction and intraocular pressure were conducted by NIDEK TONOREF II (Nidek, Gamagori, Japan).
All patients in the dataset underwent surgeries through clear corneal incision using Stellaris PC (Bausch and Lomb, Bridgewater, NJ, USA) device. Both Continuous Curvilinear Capsulorhexis (CCC) and IOL implantation in the capsular bag were performed such that the eye was stabilized using an irrigating handpiece introduced into the eye through a side port incision. FineVision Micro F trifocal IOL (Physiol, Lüttich, Belgium) was implanted. All IOLs in the dataset were calculated using SRK/T formula (Retzlaff, Sanders & Kraff, 1990) with an optimized A constant equaled to 119.1. In exceptional cases, the optical power of IOL was corrected by the decision of the surgeon, especially in the eyes with non-standard biometric specificities but all patients’ targeted refraction was on emmetropia.

At each follow-up visit, a complete slitlamp evaluation, contactless tonometry, distance objective refraction ($\text{Rx}_{\text{post}}$), distance subjective refraction, near subjective refraction, keratometry, UDVA, CDVA, uncorrected near visual acuity (UNVA), corrected near visual acuity (CNVA) measurements were performed.

The post-operative examinations were collected after at least 25 days from surgery which is the shortest time we consider for sufficient vision recovery based on conclusions in work of Conrad-Hengerer (Conrad-Hengerer et al., 2015).

**Feature selection**

Based on the database data integrity we have selected K, ACD, AL, Age, $\text{Rx}_{\text{pre}}$ as our model input parameters. $\text{Rx}_{\text{post}}$ and optical power of implanted IOL (IOL$_{\text{Implanted}}$) was used in training target definition. Potential limitation of this selection is discussed further in discussion section.

**Data mining and optimization**

The EHR system data are stored using SQL Server (Microsoft, Redmond, USA) relational database technology. A single purpose made SQL script was designed to get an initial dataview, which was then further data mined to obtain master dataset (MD). The following inclusion and exclusion criteria were used in order to filter the data from physiologically implausible entries and non-standard surgical cases.

Following inclusion criteria were used to obtain MD:
- ACD between 1 and 5 mm
- Preoperative and postoperative UDVA > CDVA in [logMAR]
- AL between 15 and 40 mm
- Mean K between 30 and 60 D
- Patient age between 18 and 99
- Optical power of implanted IOL between 6 and 35 D

Examinations and values were excluded from the MD for each eye in case of:
- Non-standard surgical procedure or intraoperative complications or any complications affecting postoperative vision recovery
  - Surgery record contains any of strings: “ruptura”, “fenestrum”, “vitrektom”, “praskl”, “sklivec”, “prolaps”, “explant”, “sulc”, “sulk”, “rzp”, “key hole”
- Had ocular disease or any corneal pathology
  - Patient finding record contains any of strings: “otok”, “striat”, “edem”, “odchlípen”, “PEX”, “jizv”, “amoc”, “aparát”, “defect”, “degener”, “endotelpati”, “fibrin”, “guttat”, “haze”, “hemoftalm”, “hemophtalm”, “luxov”, “membrán”, “precip”, “zonul”
- Previous intraocular surgery or previous corneal refractive surgery
  - Patient diagnosis record contains any of strings: “LASIK”, “LASEK”, “PRK”, “LASER”, “RELEX”, “DMEK”, “DALK”, “PKP”
- Post-operative CDVA higher than 0.3 logMAR which, is widely considered as a driving standard limit (Visual Standards for Driving in Europe, Consensus paper, European Council of Optometry and Optics)
- Incomplete biometry and refraction measurements
- Preoperative corneal astigmatism of more than 3.0 diopters
- Incomplete EHR documentation
- The difference in AL to second eye > 1 mm

All of excluded cases, which were identified using strings comes from Czech medical terminology and indicate undesirable contraindication for our application.

All samples containing outliers for K, ACD, AL, Age, Rx\textsubscript{pre}, Rx\textsubscript{post} were excluded from a MD based on ±3 sigma rule as these can be considered as an error in measurement and inappropriate for model training (Kononenko & Kukar, 2007; Leys et al., 2013).

The principle of preparing data suitable for training is to find the ideal value for the already implanted IOL (IOL\textsubscript{Ideal}). IOL\textsubscript{Ideal} is considered to be an IOL that will not induce any residual postoperative refraction for the patient’s eye or will not deviate from the intended target refraction (for distance vision this was considered as 0 D). For finding such IOL\textsubscript{Ideal}, following information are needed:

- Optical power of IOL\textsubscript{Implanted}
- Measured residual refraction Rx\textsubscript{post}
- Interrelationship of Rx\textsubscript{post} and IOL\textsubscript{Implanted}

It is generally known that 1.0 D of IOL prediction error produces approximately 0.7 D of refractive prediction error at the spectacle plane (Wang et al., 2017a). However, this is a general assumption and since eye is a complex optical system it may not reach sufficient accuracy in all eyes. The interrelationship between Rx\textsubscript{post} and IOL\textsubscript{Implanted} thus should also consider eye biometrical parameters representative of the eye optical system, such as the eye AL and the
power of the cornea $K$. The interrelationship of these two variables was determined by reversed Eye Vergence Formula Eq. (1) (Olsen, 2007; Gatinel, 2018).

$$R_{\text{theorPost}} = \frac{V}{1000} - \frac{1}{1000} \left( \frac{K}{1000} - \frac{1}{1000} \left( \frac{1}{\frac{1336}{1000} - \frac{IOL}{1336} - \frac{AL - ELP}{1336}} \right) \right)$$

**Equation 1. Reversed Eye Vergence Formula**

$R_{\text{theorPost}}$ is calculated refraction for the eye with specific $K$ in [D], $AL$ in [mm], $V$ (vertex distance) in [mm], $IOL$ in [D] and Effective Lens Position $ELP$ in [mm] calculated using recommendations by (Retzlaff, Sanders & Kraff, 1990).

Change of refraction at spectacle plane with changing the IOL power value was computed using Eq. (2), and then the IOL$_{\text{Ideal}}$ calculation is expressed by Eq. (3)

$$R_{0.5IOL} = R_{\text{theorPost}}(IOL) - R_{\text{theorPost}}(IOL + 0.5)$$

**Equation 2. Dioptric change of refraction at spectacle plane in case of IOL value change of 0.5 [D]**

$$IOL_{\text{Ideal}} = IOL_{\text{Implanted}} + \left( \frac{R_{\text{post}}}{R_{0.5IOL}} \right) \times 0.5$$

**Equation 3. Calculation of ideal value of IOL for the specific eye**

MD was then randomly divided into Selection set and Verification set in proportion 70 to 30 %. Selection set variables were normalized using mapminmax MATLAB 2017a (MathWorks, Natick, MA, USA) routine, which maps row minimal and maximal values between -1 and 1. Every Verification set variable was cleared of samples out of the minimum and maximum range of Selection set to avoid the prediction error on non-trained data. Verification set variables were then normalized using mapminmax with the same normalization parameters.

**Data description**

Selection set (70% of MD, Table 1.) contained information about 1539 eyes (771 right eyes, 768 left eyes) of 1168 patients (540 male, 628 female).
Age failed in normality by Shapiro-Wilk ($P_{SW}$; Table 1.) but was confirmed by D'Agostino-Pearson's K2 test ($P_{DP}$; Table 1.). $Rx_{pre}$ (Fig. 2A) and IOL_{Ideal} (Fig 2B) failed in normality assessment by both normality tests.

Verification set (30% of MD, Table 2.) contained information about 655 eyes (340 right eyes, 315 left eyes) of 591 patients (272 male, 319 female). As in Selection set case, only $Rx_{pre}$ (Fig. 2C) and IOL_{Ideal} (Fig. 2D) failed in normality assessment by both normality tests ($P_{SW}$, $P_{DP}$; Table 2.).

**Machine learning**

For finding the design and training of each model Selection set was used. Verification set was used for results evaluation. No samples from the Verification set were introduced to the model during the design and training phase and vice versa no samples from Selection set were used for model evaluation. Our model predictors are variables mentioned in the Feature selection section: K, ACD, AL, Age, $Rx_{pre}$. The training target was IOL_{Ideal} and the prediction outcome was IOL_{Predicted}.

**SVM - RM**

Our SVM-RM model was designed and trained in MATLAB 2017a (MathWorks, Natick, MA, USA) using fitrsvm (MathWorks, 2017a) method. Finding the appropriate hyperparameters for a given task is one of the most important steps in designing the model, on which the training and testing time but above all the accuracy of the model depends (Wang & Gong, 2018). The optimal hyperparameters of the model were found through the OptimizeHyperparameters (MathWorks, 2017a) method that searched for optimal kernel function, kernel scale, epsilon, box constraint and polynomial order.

Selection set was used in model training by Sequential minimal optimization (SMO) algorithm (Zhi-Qiang Zeng et al., 2008) with 30% of randomly selected data for holdout validation. The model parameters are summarized in Table 3.

**MLNN - EM**

For the MLNN performance improvement ensemble median was used as a better alternative to ensemble averaging reported by (Kourentzes, Barrow & Crone, 2014).

Our MLNN presented by Fig. 3. was designed and trained in MATLAB 2017a (MathWorks, Natick, MA, USA) by fitnet (MathWorks, 2017b) and had 1 hidden layer with 5 neurons and one output layer with one neuron with Linear transfer function. The internal structure and links of MLNN are described e.g. by Tuckova or in MATLAB 2017a documentation (Tuckova, 2009; MathWorks, 2017b). Hyperbolic Tangent Sigmoid transfer function was used for hidden layers and Linear transfer function for output layer.
function was used as a transfer function in hidden layer and is proposed by many authors as a
good choice for multivariate functions approximation (Anastassiou, 2011; Romero Reyes et al.,
2013). Levenberg-Marquardt backpropagation algorithm was used for model training using
trainlm (MathWorks, 2017c) method (Ranganathan, 2004).

Ensemble median factor was set to 10 which means that 10 MLNN models were trained
by Selection set in order to produce a desired output taken as a median of all outputs. Weights
and biases were initialized by Nguyen-Widrow initialization function for each ensemble training
cycle (Nguyen & Widrow, 1990).

Early stopping algorithm was used to overcome model overfitting each ensemble training
cycle. Selection set was randomly divided into three groups. For network training, validation and
testing by 70:15:15 ratio (Ross et al., 2009). MLNN training was stopped when the network
performance on the validation group failed to improve or remained the same for 20 epochs. The
weights and biases at the minimum of the validation error were returned for each ensemble
model. Training, validation and test performances for our MLNN-EM are summarized in Table
4.

The optimal number of neurons in MLNN hidden layer was found iteratively, testing all
available combinations of neurons from 1 up to 350 neurons in hidden layer. The topology
ensemble which ensured the smallest median + 1x standard deviation of the test Mean Square
Error (MSE) was selected for next processing. With the rising number of the neurons in the
hidden layer, test MSE grew also (Fig.4).

Unless otherwise mentioned, the default values of the MATLAB functions were used. All
these parameters can be found in the MATLAB documentation (MathWorks, 2019).

Evaluation methodology and statistical tests

The results predicted by each model were compared against the achieved Clinical results
(CR), and the both models were compared each to other. In the results evaluation and statistical
analysis, we followed the recommendations described in the work of Wang (Wang et al., 2017a).
The mean numerical prediction error (ME), mean absolute prediction error (MAE), median
absolute prediction error (MedAE), standard deviation (STD), Minimum prediction error (Min),
Maximum prediction error (Max) as well as percentages of eyes within prediction error (PE)
targets of ±0.25 D, ±0.50 D, ±0.75 D, ±1.00 D, were determined for Rx_{post} and refraction
calculated from IOL_{predicted} (Rx_{predicted}). Rx_{predicted} calculation describe Eq. (4).

\[
Rx_{predicted} = \left(\frac{IOL_{implanted} - IOL_{predicted}}{0.5}\right) \times RJ_{05} + Rx_{post}
\]

Equation 4. Calculation of Rx_{predicted} from IOL_{predicted}
Since AL is referred as the most important in predicting IOL power (Mahdavi & Holladay, 2011), the evaluation process is usually divided into subgroups based on AL (Wang et al., 2017a). Verification set was thus divided into the following AL subgroups:

- SHORT eyes group – eyes with AL <= 22 mm – 81 samples
- MEDIUM eyes group – eyes with 22 mm < AL < 24 mm – 480 samples
- LONG eyes group – eyes with AL => 24 mm – 94 samples
- ALL eyes group – whole Verification set with all eyes – 655 samples

Statistical analysis was performed using MATLAB 2017a (MathWorks, Natick, MA, USA). Wilcoxon test (Mercier et al., 2015) was used to assess MAE and MedAE difference between the real clinical calculation results and both models mutually. McNemar test with Yates' correction (Westfall, Troendle & Pennello, 2010) were used to evaluate the difference in the percentage of eyes in certain PE diopter group between real clinical calculation results and both models. And for both models mutually McNemar test was added by and Sing test (Dixon & Mood, 1946). Bonferroni correction was applied for multiple comparisons. The level of significance was set at 0.05 and all P values were reported.

### Results

Table 5. shows results for all evaluated parameters in ALL axial length sample group. In comparison to CR, both models showed significantly lower absolute error (SVM-RM $P=3.422\times10^{-78}$ and MLNN-EM $P=2.841\times10^{-76}$). MLNN-EM had a lower absolute error than SVM-RM but this was not statistically significant. The overall percentage of eyes with prediction errors between ±0.25 D, ±0.50 D, ±0.75 D and ±1.00 D compared to CR was significantly higher for both models (SVM-RM $P_{\pm0.25}=7.860\times10^{-7}$, $P_{\pm0.50}=0$, $P_{\pm0.75}=1.443\times10^{-15}$, $P_{\pm1.00}=4.823\times10^{-7}$ and MLNN-EM $P_{\pm0.25}=2.140\times10^{-7}$, $P_{\pm0.50}=0$, $P_{\pm0.75}=1.110\times10^{-16}$, $P_{\pm1.00}=2.992\times10^{-7}$). MLNN-EM performs better than SVM-RM in ±0.25 D, ±0.75 D and worse or same for ±0.50 D, ±1.00 D PE groups but this was not statistically significant.

Table 6. shows results for all evaluated parameters in SHORT axial length sample group. Compared to CR, both models had significantly lower absolute error (SVM-RM $P=3.674\times10^{-7}$ and MLNN-EM $P=7.445\times10^{-8}$), SVM-RM performed significantly better for ±0.50 D and ±1.00 D PE groups ($P_{\pm0.50}=0.029$ and $P_{\pm1.00}=0.041$) and better for ±0.25 D and ±0.75 D PE groups ($P_{\pm0.75}=0.070$) but this was not statistically significant, MLNN-EM performed significantly better for ±0.50 D and ±0.75 D ($P_{\pm0.50}=0.046$ and $P_{\pm0.75}=0.027$) and worse for ±0.25 D and ±1.00 D PE groups ($P_{\pm0.25}=0.429$ and $P_{\pm1.00}=0.131$) but this was not statistically significant. MLNN-EM had a lower absolute error than SVM-RM but this was not significant. MLNN-EM performs better than SVM-RM in ±0.25 D and ±0.75 D PE groups and worse or same for ±0.50 D and ±1.00 D PE groups but this was not statistically significant.
Table 7. shows results for all evaluated parameters in MEDIUM axial length sample group. Compared to CR, both models had significantly lower absolute error (SVM-RM $P=3.674e-7$ and MLNN-EM $P=7.445e-8$), both SVM-RM and MLNN-EM performed significantly better for all PE groups (SVM-RM $P_{\pm0.25}=5.699e-6$, $P_{\pm0.50}=0$, $P_{\pm0.75}=1.257e-10$, $P_{\pm1.00}=1.009e-3$ and MLNN-EM $P_{\pm0.25}=3.595e-6$, $P_{\pm0.50}=0$, $P_{\pm0.75}=2.025e-11$, $P_{\pm1.00}=3.164e-4$). MLNN-EM had a lower absolute error than SVM-RM but this was not significant. MLNN-EM performs better than SVM-RM in $\pm0.75$ D and $\pm1.0$ D PE groups and worse for $\pm0.25$ D and $\pm0.50$ D PE groups but this was not statistically significant.

Table 8. shows results for all evaluated parameters in LONG axial length sample group. Compared to CR, both models had significantly lower absolute error (SVM-RM $P=3.954e-13$ and MLNN-EM $P=1.289e-13$), both SVM-RM and MLNN-EM performed significantly better for all PE groups (SVM-RM $P_{\pm0.25}=0.041$, $P_{\pm0.50}=4.785e-5$, $P_{\pm0.75}=2.152e-5$, $P_{\pm1.00}=3.283e-3$ and MLNN-EM $P_{\pm0.25}=0.030$, $P_{\pm0.50}=4.976e-5$, $P_{\pm0.75}=2.151e-5$, $P_{\pm1.00}=3.283e-3$). MLNN-EM had a lower absolute error than SVM-RM but this was not significant. MLNN-EM performs better than SVM-RM in $\pm0.25$ D and $\pm0.50$ D PE groups but this was not statistically significant and same for $\pm0.75$ D and $\pm1.00$ D PE groups.

$P$ values for mutual evaluation of both models are presented in Table 9. For clarity, chart comparing PE of all groups is presented at Fig 5.

**Discussion**

We have described methodology of selecting and optimizing dataset for SVM-RM and MLNN-EM training, and compared accuracy of both models with the current calculation method used in the clinical practice. Overall, the percentages of eyes with prediction errors between $\pm0.25$ D, $\pm0.50$ D, $\pm0.75$ D and $\pm1.00$ D for both models were significantly better for the vast majority of evaluated parameters compared to CR. Insignificant improvement occurred only in PE $\pm0.25$ D and $\pm0.75$ D groups for the SHORT axial length subset. Anyway, as previously mentioned, calculations for eyes with a short axial length are more problematic due to the more complex ELP prediction and because of the higher probability of a steep cornea and a shallow ACD (Hoffer, 1980). Compared to CR, both models in all AL subgroups had smaller SD, which expresses higher certainty of the calculation method (Shajari et al., 2018). Long eyes over 26.3 mm and extreme long eyes were not included in this study.

Compared to the results of the Barrett Universal II formula obtained from the literature (Table 10.), which is often presented as the most accurate calculation formula, the accuracy achieved by SVM-RM and MLNN-EM is competitive (Cooke & Cooke, 2016; Kane et al., 2016, 2017; Shajari et al., 2018). For the $\pm0.50$ D PE category, the results achieved with SVM-RM and MLNN-EM were even slightly better. However, in order to objectively compare the results,
it would be necessary to evaluate all methods on the same datasets and not using the outcomes
source in the literature.

Mutual evaluation did not show a significant difference between tested models, so it can
be said that both provide similar accuracy of the calculations in all tested PE groups.

Both models predicted almost identical, and compared to the CR a slightly larger,
minimal error, which we were not expecting.

Undoubtedly, the reason for the significantly worse results of CR group is its simplicity,
where only AL and K is used for IOL power calculation. Modern calculation methods, in order
to increase the calculation accuracy, take into account more circumstances, which could affect
the refractive predictability of the surgery (Olsen, 2007; Haigis, 2012; Gökce et al., 2018). Input
parameters used in our models are standard parameters acquired using regular patient
examination prior the cataract surgery. Thus, it does not introduce any additional requirement on
the data acquisition.

Worse CR group results could be also because the non-optimized constant of implanted
IOL. This is seen in the mean error of the CR group, which has a range between -0.369 to -0.535
D among all axial length subsets. Our method of IOL\textsubscript{Ideal} calculation optimizes the mean error of
prediction to zero. This mechanism of IOL\textsubscript{Ideal} calculation can thus influence the mean error
based on the desired refraction.

Table 11. describes the input parameters used by contemporary formulas (Olsen, 2007).
Our model input parameters are K, ACD, AL, Age and Rx\textsubscript{pre}, which are the all possible
calculation variables which could be gained during the data mining process.

IOL Master 500 used in the biometry examination to gather the anatomical data is not
able to measure lens-thickness (LT). However the influence on the precision could probably be
neglected, as it is said to be the second least important calculation factor (Gale et al., 2009)
conversely, that it can have a greater influence on the IOL calculation than K (Olsen, 2006 ). One
of the other ways how to improve the accuracy of calculations would be to find a way how to
extract information from incomplete WTW measurements as this value is referred as the third
most important in predicting ELP (Mahdavi & Holladay, 2011). It is possible to find a way how
to handle missing values in datasets in order to maximize information gain (Kaiser, 2014).

In order to avoid distortion of statistical analysis by correlated data, it is recommended to
include only one eye per patient in analyses (Armstrong, 2013). Our verification set contained
less than 10% of the data that came from both eyes of the patients. This means that the intra-class
correlation factor will be less than 0.1 in the worst possible scenario (between eyes correlation
equals to 1 – for every applicable patient in Verification set) indicating extremely poor
correlation (Cicchetti, 1994; Koo & Li, 2016). We have thus concluded that it is save to use
conventional methods of statistical analysis while including maximum eyes in our datasets.
Our method does not use A constants, as usual formulas, so both models are designed as lens-specific, so the ELP prediction is coded directly into the model internal structures. The IOL power calculation for another IOL would require to go through the whole Data Preparation, Model Design and Training and Evaluation process. However, due to the fact that there are many small datasets machine learning strategies, it would not be necessary to search for such amount of training data (Jiang, Li & Zhou, 2009; Olson, Wyner & Berk, 2018). Last limitation could be the unknown training accuracy outside the input variables training range.

Conclusions

This study indicated that SVM-RM and MLNN-EM have a strong potential of improving clinical IOL calculations. Greater optimization and accuracy of IOL calculations reduces the risk of subsequent reoperation or potential refractive laser corrections and the associated risk of complications and increases patients’ comfort.

Further direction of our research and work will be directed to testing the next machine learning algorithms that might be suitable for IOL calculations such as convolutional neural networks, which are mainly used in image processing, but often in the field of biomedical engineering (Le, Ho & Ou, 2017, 2018; Le & Nguyen, 2019) and to the implementation of both models to our EHR system.

Acknowledgements

This work is a summary of independent research. The authors would like to thank the reviewers for their valuable suggestions.

References

Abell RG, Vote BJ. 2014. Cost-Effectiveness of Femtosecond Laser-Assisted Cataract Surgery versus Phacoemulsification Cataract Surgery. *Ophthalmology* 121:10–16. DOI: 10.1016/j.ophtha.2013.07.056.

Abulafia A, Barrett GD, Rotenberg M, Kleinmann G, Levy A, Reitblat O, Koch DD, Wang L, Assia EI. 2015. Intraocular lens power calculation for eyes with an axial length greater than 26.0 mm: Comparison of formulas and methods. *Journal of Cataract & Refractive Surgery* 41:548–556. DOI: 10.1016/j.jcrs.2014.06.033.

Anastassiou GA. 2011. Multivariate hyperbolic tangent neural network approximation. *Computers and Mathematics with Applications*. DOI: 10.1016/j.camwa.2010.12.029.

Armstrong RA. 2013. Statistical guidelines for the analysis of data obtained from one or both eyes. Ophthalmic and Physiological Optics 33:7–14. DOI: 10.1111/opo.12009.

Astbury N, Ramamurthy B. 2006. How to avoid mistakes in biometry. *Community Eye Health Journal*.

Chen Y-A, Hrimschall N, Findl O. 2011. Evaluation of 2 new optical biometry devices and comparison with the current gold standard biometer. *Journal of Cataract & Refractive Surgery*
Surgery 37:513–517. DOI: 10.1016/j.jcrs.2010.10.041.
Chun-Hsin Wu, Chia-Chen Wei, Da-Chun Su, Ming-Hua Chang, Jan-Ming Ho. 2003. Travel
time prediction with support vector regression. In: Proceedings of the 2003 IEEE
International Conference on Intelligent Transportation Systems. IEEE, 1438–1442. DOI:
10.1109/ITSC.2003.1252721.

Cicchetti D V. 1994. Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and
Standardized Assessment Instruments in Psychology. Psychological Assessment. DOI:
10.1037/1040-3590.6.4.284.

Clarke GP, Burmeister J. 1997. Comparison of intraocular lens computations using a neural
network versus the Holladay formula. Journal of Cataract & Refractive Surgery 23:1585–
1589. DOI: 10.1016/S0886-3350(97)80034-X.

Conrad-Hengerer I, Al Sheikh M, Hengerer FH, Schultz T, Dick HB. 2015. Comparison of
visual recovery and refractive stability between femtosecond laser–assisted cataract surgery
and standard phacoemulsification: Six-month follow-up. Journal of Cataract & Refractive
Surgery 41:1356–1364. DOI: 10.1016/j.jcrs.2014.10.044.

Cooke DL, Cooke TL. 2016. Comparison of 9 intraocular lens power calculation formulas.
Journal of Cataract & Refractive Surgery 42:1157–1164. DOI: 10.1016/j.jcrs.2016.06.029.

Dixon WJ, Mood AM. 1946. The Statistical Sign Test. Journal of the American Statistical
Association 41:557–566. DOI: 10.1080/01621459.1946.10501898.

Ferrari S, Stengel RF. 2005. Smooth Function Approximation Using Neural Networks. IEEE
Transactions on Neural Networks 16:24–38. DOI: 10.1109/TNN.2004.836233.

Frampton G, Harris P, Cooper K, Lottery A, Shepherd J. 2014. The clinical effectiveness and
cost-effectiveness of second-eye cataract surgery: a systematic review and economic
evaluation. Health Technology Assessment 18:1–206. DOI: 10.3310/hta18680.

Gale RP, Saldana M, Johnston RL, Zuberbuhler B, McKibbin M. 2009. Benchmark standards for
refractive outcomes after NHS cataract surgery. Eye. DOI: 10.1038/sj.eye.6702954.

Gatinel D. 2018. CALCULATION OF IMPLANT - THEORETICAL FORMULA. Available at
https://www.gatinel.com/en/recherche-formation/biometrie-oculaire-calcu-
dimplant/calcul-dimplant-formule-theorique/(accessed February 13, 2019).

Girosi F. 1992. Some extensions of radial basis functions and their applications in artificial
intelligence. Computers & Mathematics with Applications 24:61–80. DOI: 10.1016/0898-
1221(92)90172-E.

Gökce SE, Montes De Oca I, Cooke DL, Wang L, Koch DD, Al-Mohtaseb Z. 2018. Accuracy of
8 intraocular lens calculation formulas in relation to anterior chamber depth in patients with
normal axial lengths. Journal of Cataract & Refractive Surgery 44:362–368. DOI:
10.1016/j.jcrs.2018.01.015.

Haag-Streit AG Koeniz S. 2017. Hill-RBF Method - White Paper.

Haigis W. 2012. Challenges and approaches in modern biometry and IOL calculation. Saudi
Journal of Ophthalmology. DOI: 10.1016/j.sjopt.2011.11.007.

Herbrich R. 1999. Support vector learning for ordinal regression. In: 9th International
Conference on Artificial Neural Networks: ICANN ’99. IEE, 97–102. DOI:
10.1049/cp:19991091.

Hill W. Hill-RBF Calculator. Available at www.rbfcalculator.com (accessed December 9, 2018).

Hoffer KJ. 1980. Biometry of 7,500 cataractous eyes. American Journal of Ophthalmology. DOI:
10.1016/S0002-9394(14)74917-7.

Jap D, Stöttinger M, Bhasin S. 2015. Support vector regression. In: Proceedings of the Fourth
503  Workshop on Hardware and Architectural Support for Security and Privacy - HASP ’15.
504  DOI: 10.1145/2768566.2768568.
505  Jiang Y, Li M, Zhou ZH. 2009. Mining extremely small data sets with application to software
506  reuse. Software - Practice and Experience. DOI: 10.1002/spe.905.
507  Kaiser J. 2014. Dealing with Missing Values in Data. Journal of Systems Integration. DOI:
508  10.20470/jsi.v5i1.178.
509  Kane JX, Van Heerden A, Atik A, Petsoglou C. 2016. Intraocular lens power formula accuracy:
510  Comparison of 7 formulas. Journal of Cataract & Refractive Surgery 42:1490–1500. DOI:
511  10.1016/j.jcrs.2016.07.021.
512  Kane JX, Van Heerden A, Atik A, Petsoglou C. 2017. Accuracy of 3 new methods for
513  intraocular lens power selection. Journal of Cataract & Refractive Surgery 43:333–339.
514  DOI: 10.1016/j.jcrs.2016.12.021.
515  Koch DD, Hill W, Abulafia A, Wang L. 2017. Pursuing perfection in intraocular lens
516  calculations: I. Logical approach for classifying IOL calculation formulas. Journal of
517  Cataract and Refractive Surgery. DOI: 10.1016/j.jcrs.2017.06.006.
518  Kononenko I, Kukar M. 2007. Machine learning and data mining. Woodhead Publishing
519  Limited. DOI: 10.1533/9780857099440.
520  Koo TK, Li MY. 2016. A Guideline of Selecting and Reporting Intraclass Correlation
521  Coefficients for Reliability Research. Journal of chiropractic medicine. DOI:
522  10.1016/j.jcm.2016.02.012.
523  Kourentzes N, Barrow DK, Crone SF. 2014. Neural network ensemble operators for time series
524  forecasting. Expert Systems with Applications. DOI: 10.1016/j.eswa.2013.12.011.
525  Kurban T, Beşdok E. 2009. A Comparison of RBF Neural Network Training Algorithms for
526  Inertial Sensor Based Terrain Classification. Sensors 9:6312–6329. DOI:
527  10.3390/s90806312.
528  Lampariello F, Sciandrone M. 2001. Efficient training of RBF neural networks for pattern
529  recognition. IEEE Transactions on Neural Networks 12:1235–1242. DOI:
530  10.1109/72.950152.
531  Lee TH, Sung MS, Cui L, Li Y, Yoon KC. 2015. Factors Affecting the Accuracy of Intraocular
532  Lens Power Calculation with Lenstar. Chonnam Medical Journal 51:91. DOI:
533  10.4068/cmj.2015.51.2.91.
534  Le N-Q-K, Ho Q-T, Ou Y-Y. 2017. Incorporating deep learning with convolutional neural
535  networks and position specific scoring matrices for identifying electron transport proteins.
536  Journal of Computational Chemistry 38:2000–2006. DOI: 10.1002/jcc.24842.
537  Le N-Q-K, Ho Q-T, Ou Y-Y. 2018. Classifying the molecular functions of Rab GTPases in
538  membrane trafficking using deep convolutional neural networks. Analytical Biochemistry
539  555:33–41. DOI: 10.1016/j.ab.2018.06.011.
540  Le NQK, Nguyen V-N. 2019. SNARE-CNN: a 2D convolutional neural network architecture to
541  identify SNARE proteins from high-throughput sequencing data. PeerJ Computer Science.
542  DOI: 10.7717/peerj-cs.177.
543  Le NQK, Ou YY. 2016a. Incorporating efficient radial basis function networks and significant
544  amino acid pairs for predicting GTP binding sites in transport proteins. BMC
545  Bioinformatics. DOI: 10.1186/s12859-016-1369-y.
546  Le NQK, Ou YY. 2016b. Prediction of FAD binding sites in electron transport proteins
547  according to efficient radial basis function networks and significant amino acid pairs. BMC
548  Bioinformatics. DOI: 10.1186/s12859-016-1163-x.
Le NQK, Yapp EKY, Ho Q-T, Nagasundaram N, Ou Y-Y, Yeh H-Y. 2019. iEnhancer-5Step: Identifying enhancers using hidden information of DNA sequences via Chou’s 5-step rule and word embedding. Analytical Biochemistry 571:53–61. DOI: 10.1016/j.ab.2019.02.017.

Leys C, Ley C, Klein O, Bernard P, Licata L. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology. DOI: 10.1016/j.jesp.2013.03.013.

Linebarger EJ, Hardten DR, Shah GK, Lindstrom RL. 1999. Phacoemulsification and modern cataract surgery. Survey of Ophthalmology. DOI: 10.1016/S0039-6257(99)00085-5.

Mahdavi S, Holladay J. 2011. IOLMaster® 500 and Integration of the Holladay 2 Formula for Intraocular Lens Calculations. European Ophthalmic Review 05:134. DOI: 10.17925/EOR.2011.05.02.134.

MathWorks. 2017a. Fit a support vector machine regression model.

MathWorks. 2017b. Function fitting neural network - MATLAB fitnet - MathWorks Benelux.

MathWorks. 2017c. Levenberg-Marquardt backpropagation.

MathWorks. Matlab documentation, 2019. Available at https://www.mathworks.com/help/

Melles RB, Holladay JT, Chang WJ. 2018. Accuracy of Intraocular Lens Calculation Formulas. Ophthalmology 125:169–178. DOI: 10.1016/j.ophtha.2017.08.027.

Mercier É, Even A, Mirisola E, Wolfersberger D, Sciamanna M. 2015. Numerical study of extreme events in a laser diode with phase-conjugate optical feedback. Physical Review E 91:042914. DOI: 10.1103/PhysRevE.91.042914.

Mongillo M. 2011. Choosing Basis Functions and Shape Parameters for Radial Basis Function Methods. SIAM Undergraduate Research Online 4:190–209. DOI: 10.1137/11S010840.

Nguyen D, Widrow B. 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In: 1990 IJCNN International Joint Conference on Neural Networks. DOI: 10.1109/IJCNN.1990.137819.

Norrby S. 2008. Sources of error in intraocular lens power calculation. Journal of Cataract & Refractive Surgery 34:368–376. DOI: 10.1016/j.jcrs.2007.10.031.

Olsen T. 2007. Calculation of intraocular lens power: A review. Acta Ophthalmologica Scandinavica. DOI: 10.1111/j.1755-3768.2007.00879.x.

Olson M, Wyner A, Berk R. 2018. Modern Neural Networks Generalize on Small Data Sets. In: Neural Information Processing Systems (NIPS).

Park J, Sandberg IW. 1991. Universal Approximation Using Radial-Basis-Function Networks. Neural Computation 3:246–257. DOI: 10.1162/neco.1991.3.2.246.

Pascolini D, Mariotti SP. 2012. Global estimates of visual impairment: 2010. British Journal of Ophthalmology 96:614–618. DOI: 10.1136/bjophthalmol-2011-300539.

Ranganathan A. 2004. The Levenberg-Marquardt Algorithm. Tutorial on LM algorithm. DOI: http://dx.doi.org/10.1.1.10.2258.

Retzlaff JA, Sanders DR, Kraff MC. 1990. Development of the SRK/T intraocular lens implant power calculation formula. Journal of Cataract & Refractive Surgery 16:333–340. DOI: 10.1016/S0886-3350(13)80705-5.

Roberts T V., Hodge C, Sutton G, Lawless M. 2018. Comparison of Hill-radial basis function, Barrett Universal and current third generation formulas for the calculation of intraocular lens power during cataract surgery. Clinical & Experimental Ophthalmology 46:240–246. DOI: 10.1111/ceo.13034.

Romero Reyes I V., Fedyushkina I V., Skvortsov VS, Filimonov DA. 2013. Prediction of progesterone receptor inhibition by high-performance neural network algorithm.
Ross KA, Jensen CS, Snodgrass R, Dyreson CE, Jensen CS, Snodgrass R, Skiadopoulos S, Sirangelo C, Larsgaard ML, Grahne G, Kifer D, Jacobsen H-A, Hinterberger H, Deutsch A, Nash A, Wada K, Aalst WMP, Dyreson C, Mitra P, Witten IH, Liu B, Aggarwal CC, Özsu MT, Ogbuji C, Patel C, Weng C, Patel C, Weng C, Wright A, Shabo (Shvo) A, Russler D, Rocha RA, Russler D, Lussier YA, Chen JL, Russler D, Zaki MJ, Corral A, Vazirgiannis M, Davidson I, Sarawagi S, Peyton L, Hinterberger H, Speege G, Vianu V, Gucht D Van, Etzion O, Etzion O, Curbera F, Ericsson A, Berndtsson M, Mellin J, Aalst WMP, Gray PMD, Trajcevski G, Wolfson O, Scheuermann P, Dorai C, Weiner M, Borgida A, Mylopoulos J, Vossen G, Reuter A, Grahne G, Tannen V, Elnikety S, Fekete A, Bertossi L, Geerts F, Geerts F, Fan W, Westerveld T, Jacobsen H-A, Gurrin C, Westerveld T, Etzion O, Kekäläinen J, Arvola P, Junkkari M, Wada K, Mouratidis K, Yu JX, Yao Y, Gehrke J, Babu S, Reuter A, Palmer N, Leung CK-S, Aalst WMP, Carroll MW, Gokhale A, Ouzanni M, Medjahed B, Elmagarmid AK, Manegold S, Cormode G, Mankovskii S, Zhang D, Härder T, Gao W, Niu C, Li Q, Yang Y, Refaeilzadeh P, Tang L, Liu H, Pedersen TB, Morfonios K, Ioannidis Y, Böhlen MH, Jensen CS, Snodgrass RT, Chen L. 2009. Cross-Validation. In: Encyclopedia of Database Systems. Boston, MA: Springer US, 532–538. DOI: 10.1007/978-0-387-39940-9_565.

Shajari M, Kolb CM, Petermann K, Böhm M, Herzog M, De’Lorenzo N, Schönbrunn S, Kohnen T. 2018. Comparison of 9 modern intraocular lens power calculation formulas for a quadrifocal intraocular lens. Journal of Cataract & Refractive Surgery 44:942–948. DOI: 10.1016/j.jcrs.2018.05.021.

Shammas HJ, Shammas MC. 2015. Measuring the cataractous lens. Journal of Cataract & Refractive Surgery 41:1875–1879. DOI: 10.1016/j.jcrs.2015.10.036.

Shrivastava AK, Behera P, Kumar B, Nanda S. 2018. Precision of intraocular lens power prediction in eyes shorter than 22 mm: An analysis of 6 formulas. Journal of Cataract & Refractive Surgery 44:1317–1320. DOI: 10.1016/j.jcrs.2018.07.023.

Smola AJ, Schölkopf B. 2004. A tutorial on support vector regression. Statistics and Computing 14:199–222. DOI: 10.1023/B:STCO.0000035301.49549.88.

Snyder, E. M. Hill-RBF Calculator in Clinical Practice. Available at https://crstodayeurope.com/articles/new-frontiers-in-iol-prediction-for-improved-refractive-outcomes/hill-rbf-calculator-in-clinical-practice/ (accessed January 22, 2019).

The American Society of Cataract and Refractive Surgery. ASCRS Announces Hill-RBF Calculator for Cataract Surgery IOL Power Calculations. Available at http://ascrs.org/about-ascrs/news-about/ascrs-announces-hill-rbf-calculator-cataract-surgery-iol-power-calculations (accessed December 5, 2018).

Thulasi P, Khandelwal SS, Bradley Randleman J. 2016. Intraocular lens alignment methods. Current Opinion in Ophthalmology. DOI: 10.1097/ICO.0000000000000225.

Trafalis TB, Ince H. 2000. Support vector machine for regression and applications to financial forecasting. In: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium. DOI: 10.1109/IJCNN.2000.859420.

Tuckova J. 2009. Selected applications of the artificial neural networks at the signal processing. Prague: Nakladatelstvi ČVUT.

Wang B, Gong NZ. 2018. Stealing Hyperparameters in Machine Learning. In: 2018 IEEE
Symposium on Security and Privacy (SP). IEEE, 36–52. DOI: 10.1109/SP.2018.00038.

Wang L, Koch DD, Hill W, Abulafia A. 2017a. Pursuing perfection in intraocular lens calculations: III. Criteria for analyzing outcomes. *Journal of Cataract & Refractive Surgery* 43:999–1002. DOI: 10.1016/j.jcrs.2017.08.003.

Wang W, Yan W, Fotis K, Prasad NM, Lansingh VC, Taylor HR, Finger RP, Facciolo D, He M. 2017b. Cataract Surgical Rate and Socioeconomics: A Global Study. *Investigative Ophthalmology & Visual Science* 57:5872. DOI: 10.1167/iovs.16-19894.

Westfall PH, Troendle JF, Pennello G. 2010. Multiple McNemar Tests. *Biometrics* 66:1185–1191. DOI: 10.1111/j.1541-0420.2010.01408.x.

Wu Y, Wang H, Zhang B, Du K-L. 2012. Using Radial Basis Function Networks for Function Approximation and Classification. *ISRN Applied Mathematics* 2012:1–34. DOI: 10.5402/2012/324194.

Yamaguchi T, Negishi K, Tsubota K. 2011. Functional visual acuity measurement in cataract and intraocular lens implantation. *Current Opinion in Ophthalmology*. DOI: 10.1097/ICU.0b013e3283414f36.

Yu PS, Chen ST, Chang IF. 2006. Support vector regression for real-time flood stage forecasting. *Journal of Hydrology*. DOI: 10.1016/j.jhydrol.2006.01.021.

Zhi-Qiang Zeng, Hong-Bin Yu, Hua-Rong Xu, Yan-Qi Xie, Ji Gao. 2008. Fast training Support Vector Machines using parallel sequential minimal optimization. In: *2008 3rd International Conference on Intelligent System and Knowledge Engineering*. IEEE, 997–1001. DOI: 10.1109/ISKE.2008.4731075.
Table 1 (on next page)

Selection set population characteristics

Standard Deviation (Std), Minimum (Min), Maximum (Max), Shapiro-Wilk P value (pSW) and D'Agostino-Pearson's K2 P value (pDP). Selection set was assessed for normality by Shapiro-Wilk and D'Agostino-Pearson's K2 normality tests at level of \( P = 0.001 \).
|                  | Mean  | Median | Std  | Min   | Max   | P_{SW}       | P_{DP}    |
|------------------|-------|--------|------|-------|-------|--------------|----------|
| Age [years]      | 56.89 | 57.00  | 7.25 | 36.00 | 78.00 | 8.543e-5     | 0.091    |
| K [D]            | 43.27 | 43.25  | 1.40 | 39.39 | 47.51 | 0.252        | 0.547    |
| ACD [mm]         | 3.10  | 3.10   | 0.32 | 2.21  | 4.10  | 0.189        | 0.350    |
| AL [mm]          | 23.03 | 23.07  | 0.92 | 19.94 | 26.26 | 0.010        | 0.111    |
| Rx_{pre} [D]     | 1.85  | 1.88   | 1.52 | -3.88 | 6.63  | 0.000        | 0.000    |
| IOL_{Ideal} [D]  | 22.80 | 22.50  | 2.74 | 12.62 | 34.17 | 8.615e-12    | 9.992e-16|


Table 2 (on next page)

Verification set population characteristics

Standard Deviation (Std), Minimum (Min), Maximum (Max), Shapiro-Wilk P value (pSW) and D'Agostino-Pearson's K2 P value (pDP)
|                          | Mean | Median | Std  | Min  | Max  | P<sub>SW</sub> | P<sub>DP</sub> |
|--------------------------|------|--------|------|------|------|---------------|---------------|
| Age [years]              | 56.83| 56.00  | 7.29 | 37.00| 76.00| 0.003         | 0.161         |
| K [D]                    | 43.33| 43.30  | 1.33 | 39.41| 46.92| 0.263         | 0.199         |
| ACD [mm]                 | 3.11 | 3.10   | 0.32 | 2.29 | 4.06 | 0.183         | 0.206         |
| AL [mm]                  | 23.03| 22.99  | 0.90 | 20.17| 25.88| 0.530         | 0.417         |
| Rx<sub>pre</sub> [D]     | 1.83 | 1.75   | 1.49 | -3.88| 6.63 | 1.998e-15     | 0             |
| IOL<sub>Ideal</sub> [D]  | 22.71| 22.42  | 2.64 | 15.32| 33.51| 7.793e-7      | 3.467e-7      |
**Table 3** (on next page)

SVM-RM parameters

MSE – Mean squared error
| Parameter          | Value     |
|--------------------|-----------|
| Kernel function    | Polynomial|
| Kernel Scale       | -         |
| Epsilon            | 0.0282    |
| Box constraint     | 0.0049    |
| Polynomial order   | 2         |
| MSE                | 0.0032    |
Table 4 (on next page)

MLNN-EM design parameters

MSE – Mean squared error
|            | Mean    | Median   | Std       | Min      | Max      |
|------------|---------|----------|-----------|----------|----------|
| Train MSE  | 0.00302 | 0.00306  | 9.44729E-05 | 0.0028  | 0.00311  |
| Validation MSE | 0.00307 | 0.00310  | 0.00033   | 0.0025  | 0.00364  |
| Test MSE   | 0.00329 | 0.00333  | 0.00039   | 0.0025  | 0.00387  |
| Epoch      | 22.8    | 21.5     | 18.6      | 7        | 72       |
Table 5 (on next page)

Prediction errors in the ALL axial length group for Clinical Results (CR), SVM-RM and MLNN-EM

Mean prediction error (ME), Mean absolute prediction error (MAE), Median absolute prediction error (MedAE), Standard deviation (Std), Minimum prediction error (Min), Maximum prediction error (Max), Prediction error (PE)
|                  | CR  | SVM-RM | MLNN-EM |
|------------------|-----|--------|---------|
| ME               | -0.464 | 0.012 | 0.002  |
| MAE              | 0.523 | 0.310 | 0.309  |
| MedAE            | 0.500 | 0.260 | 0.258  |
| Std              | 0.433 | 0.395 | 0.395  |
| Min              | -1.875 | -1.480 | -1.514 |
| Max              | 1.125 | 1.372 | 1.310  |
| Eyes within PE (%) |     |       |         |
| ±0.25            | 33.4 | 48.2  | 48.9   |
| ±0.50            | 57.7 | 82.8  | 82.3   |
| ±0.75            | 79.4 | 93.4  | 93.7   |
| ±1.00            | 91.8 | 97.7  | 97.7   |
Table 6 (on next page)

Prediction errors in the SHORT axial length group for Clinical Results (CR), SVM-RM and MLNN-EM

Mean prediction error (ME), Mean absolute prediction error (MAE), Median absolute prediction error (MedAE), Standard deviation (Std), Minimum prediction error (Min), Maximum prediction error (Max), Prediction error (PE)
|                      | CR   | SVM-RM | MLNN-EM |
|----------------------|------|--------|---------|
| ME                   | -0.369 | 0.002  | 0.018   |
| MAE                  | 0.465 | 0.322  | 0.320   |
| MedAE                | 0.500 | 0.302  | 0.266   |
| Std                  | 0.464 | 0.399  | 0.398   |
| Min                  | -1.500 | -0.865 | -0.930  |
| Max                  | 1.125 | 0.929  | 1.007   |
| Eyes within PE (%)   |      |        |         |
| ±0.25                | 40.7 | 44.4   | 48.1    |
| ±0.50                | 63.0 | 76.5   | 76.5    |
| ±0.75                | 85.2 | 93.8   | 95.1    |
| ±1.00                | 92.6 | 100.0  | 98.8    |
Table 7 (on next page)

Prediction errors in the MEDIUM axial length group for Clinical Results (CR), SVM-RM and MLNN-EM

Mean prediction error (ME), Mean absolute prediction error (MAE), Median absolute prediction error (MedAE), Standard deviation (Std), Minimum prediction error (Min), Maximum prediction error (Max), Prediction error (PE)
|               | CR  | SVM-RM | MLNN-EM |
|---------------|-----|--------|---------|
| ME            | -0.466 | 0.024  | 0.008   |
| MAE           | 0.523  | 0.307  | 0.307   |
| MedAE         | 0.500  | 0.251  | 0.254   |
| Std           | 0.424  | 0.396  | 0.395   |
| Min           | -1.875 | -1.480 | -1.514  |
| Max           | 0.875  | 1.372  | 1.310   |
| Eyes within PE (%) |      |        |         |
| ±0.25         | 33.1  | 49.6   | 49.4    |
| ±0.50         | 56.9  | 83.8   | 82.9    |
| ±0.75         | 79.8  | 93.3   | 93.5    |
| ±1.00         | 92.9  | 97.3   | 97.5    |
**Table 8** (on next page)

Prediction errors in the LONG axial length group for Clinical Results (CR), SVM-RM and MLNN-EM

Mean prediction error (ME), Mean absolute prediction error (MAE), Median absolute prediction error (MedAE), Standard deviation (Std), Minimum prediction error (Min), Maximum prediction error (Max), Prediction error (PE)
|                  | CR  | SVM-RM | MLNN-EM |
|------------------|-----|--------|---------|
| ME               | -0.535 | -0.043 | -0.043 |
| MAE              | 0.574  | 0.316  | 0.311   |
| MedAE            | 0.500  | 0.270  | 0.269   |
| Std              | 0.442  | 0.387  | 0.393   |
| Min              | -1.625 | -1.013 | -1.000  |
| Max              | 0.875  | 1.096  | 1.230   |
| Eyes within PE (%) |       |        |         |
| ±0.25            | 28.7  | 44.7   | 46.8    |
| ±0.50            | 57.4  | 83.0   | 84.0    |
| ±0.75            | 72.3  | 93.6   | 93.6    |
| ±1.00            | 85.1  | 97.9   | 97.9    |
Table 9 (on next page)

Mutual evaluation of difference between SVM-RM and MLNN-EM

Absolute prediction error (PE) by Wilcoxon test (WT), McNemar test (MN), Sign test (ST)
|       | ALL | SHORT | MEDIUM | LONG |
|-------|-----|-------|--------|------|
| PE WT | 0.679 | 0.763 | 0.545 | 0.917 |
| ±0.25 MN | 0.819 | 0.449 | 0.891 | 0.802 |
| ±0.50 MN | 0.735 | 0.723 | 0.540 | 1    |
| ±0.75 MN | 0.789 | 1     | 1     | 0    |
| ±1.00 MN | 0.723 | 1     | 1     | 0.479 |
| ±0.25 ST | 0.819 | 0.453 | 0.891 | 0.803 |
| ±0.50 ST | 0.735 | 1     | 0.541 | 1    |
| ±0.75 ST | 0.790 | 1     | 1     | 1    |
| ±1.00 ST | 1     | 1     | 1     | 1    |
Table 10 (on next page)

Prediction error comparison for Barrett Universal II, SVM-RM and MLNN-EM for all axial lengths

Prediction error (PE)
| Eyes within PE (%) | Barrett Universal II | SVM-RM | MLNN-EM |
|-------------------|----------------------|--------|---------|
| ±0.25             | 43.5 – 60.0          | 48.1   | 48.5    |
| ±0.50             | 72.3 – 80.6          | 82.7   | 82.3    |
| ±1.00             | 94.5 – 99.7          | 97.7   | 97.7    |
Table 11 (on next page)

Overview of contemporary formulas input parameters

K (mean keratometry), AL (axial length), ACD (anterior chamber depth), LT (lens thickness), WTW (white to white), Age (patients age), Rx-pre (preoperative refraction)
|   | Hill-RBF | HofferQ | Holladay 1 | Holladay 2 | SRK/T | Haigis | Olsen |
|---|----------|---------|------------|------------|-------|--------|-------|
| K | x        | x       | x          | x          | x     | x      | x     |
| AL| x        | x       | x          | x          | x     | x      | x     |
| ACD| x |          | x          | x          |       |        | x     |
| LT|          |         | x          | x          |       |        | x     |
| WTW| x |          | x          |            |       |        | x     |
| Age|          |         |            | x          |       |        | x     |
| Rx-pre| |          |            |            |       |        | x     |
Figure 1

Research workflow
Figure 2

Histograms

(A) $R_{x_{\text{pre}}}$ - Selection set. (B) $IOL_{\text{ideal}}$ - Selection set. (C) $R_{x_{\text{pre}}}$ - Verification set. (D) $IOL_{\text{ideal}}$ - Verification set.
Figure 3

MLNN layer structure
Figure 4

*MSE dependence on the number of neurons in the hidden layer*

*Mean Square Error (MSE)*
Figure 5

Histograms of PE in different eye AL groups

Prediction error (PE)