Multinomial Logistic Regression and Support Vector Machine for Osteoarthritis Classification

C Aroef¹, R P Yuda¹, Z Rustam¹, and J Pandelaki²

¹Department of Mathematics, University of Indonesia, Depok 16426, Indonesia
²Department of Radiology, RSUPN dr. Cipto Mangunkusumo, Jakarta 10430, Indonesia

e-mail: *chelvian.aroef@ui.ac.id, rivan.pratama@sci.ui.ac.id, rustam@ui.ac.id

Abstract. Everyone joints go through a cycle of damage and repair during their lifetime, but sometimes the body’s process to repair our joints can cause changes in their shape or structure. When these changes happen, it’s known as osteoarthritis. Osteoarthritis is the most common form of arthritis, affecting millions of people worldwide. Osteoarthritis causes pain, swelling, stiffness in the areas, and decreased the ability to move for the sufferers. Therefore it requires accurate method of classification. Many methods have been used to classify osteoarthritis, but this study will apply Multinomial Logistic Regression and Support Vector Machine (SVM) as the machine learning methods. We used CT scan result data from RSUPN dr. Cipto Mangunkusumo, Central Jakarta. The results show the SVM provides better results than Multinomial Logistic Regression in terms of classification accuracy. The highest accuracy of SVM reaches around 85%, while Multinomial Logistic Regression only 71%.

1. Introduction

All joints go through a cycle of damage and repair during their lifetime, but sometimes the process of repairing the joints may cause changes in their shape or even structure. When these changes happen in one or more of the joints in the knee, it’s known as osteoarthritis (OA) [1]. In osteoarthritis, the cartilage in the knee joint gradually wears away [2]. As the cartilage wears away, it becomes frayed and rough, and the protective space between the bones decreases [2]. As the cartilage getting rougher, the knee doesn’t move as smoothly as it should, and it might feel painful and stiff [1].

Osteoarthritis usually happens to older people because of the ability of cartilage to heal decreases [3]. As the person getting heavier, the risk of getting OA is also getting higher, because the joints will get more pressure from the body [3]. Every pound of weight gain adds 3 to 4 pounds of extra weight on the knees [3]. OA may also happen to an athlete, cause the overuse of the knee [3]. If you feel the pain when you are active, but gets a little better with the rest, maybe you should check your knee because it’s one of the symptoms of OA [3]. The other symptoms are the joints become stiff and swollen, making it difficult to bend and straighten the knee [1]. OA impacts most of your activity because it damages the knee which mostly used for activities. There are several treatment options available for OA, one of them is exercise [3]. Exercise will strengthen the muscle around the knee and makes the joint more stable and decrease pain [3].
The other option is losing some weight, losing even a small amount of weight can significantly decrease knee pain from OA.

The previous study on OA classification used AdaBoost Support Vector Machine [4]. This study will apply another method of machine learning which never been used before for Osteoarthritis classification, such as Multinomial Logistic Regression (MLR) and Support Vector Machine (SVM). Multinomial Logistic Regression is a development method of Logistic Regression when the dependent variable is more than two classes. Super Vector Machine is known for its classification method for two classes or more. Both methods are applied because the OA is going to be classified into 4 classes. We will these methods in terms of accuracy. By using both methods, it is expected to help the health sector to be able to classify osteoarthritis more efficiently.

2. Methods

2.1. Data
The data for this study is obtained from dr. RSUPN Cipto Mangunkusumo (RSCM), Central Jakarta, by the result of Cartigram scan, which composed of six features, which are: Gender, Age, Femur Lateral, Tibia Lateral, Femur Medial, and Tibia Medial, and one prediction class. The data includes 42 observations which classify into four classes.

| Number of Patient | Gender | Age | Femur Lateral | Tibia Lateral | Femur Medial | Tibia Medial | Grade OA |
|------------------|--------|-----|---------------|---------------|--------------|--------------|----------|
| 1                | M      | 50  | 45.5          | 50.5          | 85           | 36.5         | 50       |
| 2                | M      | 52  | 40.5          | 70            | 70           | 35           | 1        |
| 3                | F      | 60  | 45            | 40            | 70           | 30           | 4        |
| 4                | F      | 70  | 35            | 30            | 45           | 20           | 2        |

2.2. Logistic Regression
Logistic Regression in many ways is the natural complement of ordinary linear regression whenever the target variable is categorized.

For a target (dependent) variable $Y$ with two class and predictor (independent) variable $X$, let $g(x) = Pr (X = x) = 1 − Pr (X = x)$, the logistic regression model has a linear form for Logit with probability as follows:

$$Logit[g(x)] = \log \left( \frac{g(x)}{1 - g(x)} \right) = \alpha + \beta x,$$

where the odds $= \frac{g(x)}{1 - g(x)}$ (1)

The Logit has a form of linear approximation, and logit is equaled with the logarithm of the odds. The parameter $\beta$ is the rate of increase or decrease of the S-shaped curve of $g(x)$.

2.3. Multinomial Logistic Regression
The logistic regression can be extended to models with multiple predictor variables. Let $u$ define the number of predictors for a binary response $Y$ by $x_1, x_2, \ldots, x_u$. The model for log odds as follow:

$$Logit[Pr Pr (Y = 1)] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_u x_u$$

and the alternative formula, directly specifying $h(x)$ is

$$h(x) = \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_u x_u)}{1 - \exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_u x_u)}$$

(2)
The parameter $\beta_i$, for $1 < i \leq u$ and $1 < j \leq u$ and $i \neq j$, refers to the effect of $x_i$ on the log odds that $Y = 1$, controlling other $x_j$, for instance, $\exp(\beta_i)$ is the multiplicative effect on the odds of a one-unit increase $x_i$, at fixed level of other $x_j$.

If we have $n$ independent observations with $u$-predictor variables, and the target variable has $q$ categories, to build the logits in the multinomial case, one of the categories must be considered as the base level and all the logits functions are constructed relative to it. There are no conditions to turn the category into the base level, so we can choose randomly, we will turn category $z$ as the base level. Since there is no order, it is possible that any category could be labeled $z$. Let $H_j$ denote the multinomial probability of observation for the $j^{th}$ category, to find the relationship between this probability and the $u$-predictor variables, $X_1, X_2, ..., X_u$, the multinomial logistic regression model is:

$$
\log \frac{H_j(x)}{H_z(x)} = a_{0j} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \cdots + \beta_{uj}x_{ui}
$$

(4)

where $j = 1, 2, ..., z - 1$ and $i = 1, 2, ..., n$. Since all the $H$'s adds to unity this reduces to

$$
\log \log \left( \frac{H_j(x)}{H_z(x)} \right) = \frac{\exp(a_{0j} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \cdots + \beta_{uj}x_{ui})}{1 + \sum_{j=1}^{z-1} \exp(a_{0j} + \beta_{1j}x_{1i} + \beta_{2j}x_{2i} + \cdots + \beta_{uj}x_{ui})}
$$

(5)

for $j = 1, 2, ..., z - 1$. The model parameters are determined by the method of multinomial linear.

2.4. Support Vector Machine

Support Vector Machine (SVM) is one of a machine learning algorithms usually used for classification and regression introduced by Russian Scientist, Vapnik in 1992. For the first time of SVM appearance, SVM was used only for binary classification, but now it could be used for more than two-class classification, usually called multiclass classification. SVM takes the form of mapping input space into higher dimensional space to support nonlinear classification problems where the maximum separation of the hyperplane is constructed. The hyperplane is a linear pattern (depends on the kernel) that maximizes the margin resulting maximum value between classification classes.

2.4.1. Characteristic of SVM [6]. Given dataset \{ $x_m, y_m$\}_{m=1}^{N} where $N$ is the number of samples, $x_m \in \mathbb{R}^D$ is a feature vector from sample-$m$, where $D$ is the number of features (dimension), and $y_m$ is the class label. For binary classification $y_m \in \{-1, +1\}$, but for more than two (multiclass) classification problem $y_m \in \{1, 2, ..., u\}$ where $u$ is the number of classes. The main goal of using Support Vector Machine is to find the best hyperplane for the data:

$$
q \cdot x + p = 0
$$

(6)
The problem of SVM optimization can be summarized as follows:

\[
\begin{align*}
\frac{1}{2} \|q\|^2 \quad (7)
\end{align*}
\]

s.t. \(y_i (q^T \cdot x_i + p) \geq 1, \forall m = 1, ..., N\)  \( (8)\)

The objective function (7) to find \(q \in R^n\) and \(p \in R^n\) subject to (8), where \(q\) is the weights vector and \(p\) is bias. By solving the equation (7) and (8), we get the equation of \(q\) and \(p\) as follows:

\[
q = \sum_{m=1}^{N} a_m y_m x_m \quad (9)
\]

\[
p = \frac{1}{N} \sum_{m \in S} \left( y_m - \sum_{m \in S} a_m y_m x_m \right) \quad (10)
\]

By the equation we get the decision function as follows:

\[
f(x) = q \cdot x + p \quad (11)
\]

which could maximize the margins.

3. Experimental Result

This study used software R Studio version 1.1.463 for both Multinomial Logistic Regression and Support Vector Machine.

3.1. Osteoarthritis Classification using Multinomial Logistic Regression

Table 2. Results of Accuracy Osteoarthritis Classification using Multinomial Logistic Regression.

| Data Training (%) | Accuracy (%) |
|-------------------|--------------|
| 10                | 50           |
| 20                | 54.54        |
| 30                | 51.85        |
| 40                | 66.67        |
| 50                | 60           |
As a result of the table, the data training at 70% has the best accuracy among the rest with around 72.72%. The worst accuracy is recorded at data training 10% and 90% with 50%.

### 3.2. Osteoarthritis Classification using Support Vector Machine

#### Table 3. Results of Accuracy Osteoarthritis Classification using Support Vector Machine with Radial Basis Kernel Function, parameter $C=4$, and $\frac{1}{\sigma^2} = 0.1259338$.

| Data Training (%) | Accuracy (%) |
|-------------------|--------------|
| 10                | 67.56        |
| 20                | 54.54        |
| 30                | 62.06        |
| 40                | 66.67        |
| 50                | 60           |
| 60                | 62.50        |
| 70                | 72.72        |
| 80                | 85.71        |
| 90                | 67.67        |

As a result of the table, the data training at 80% has the best accuracy among the rest with around 85.71%. The worst accuracy is recorded at data training 20% with 54.54%.

### 4. Conclusion

This study applied two different classifiers methods, which are Multinomial Logistic Regression (MLR) and Support Vector Machine (SVM) to compare the accuracy. From the simulation for SVM with Radial Basis Function (RBF) kernel, we get the best parameter which is $C = 4$ and $\sigma = 0.1259338$. From the experiment, we get the best accuracy is around 72.72% for Multinomial Logistic Regression from 70% training data. Support Vector Machine has better accuracy for 85.71% from 80% training data. These results show the performance of Support Vector Machine with RBF kernel is better than Multinomial Logistic Regression in terms of accuracy.

### Acknowledgment

This research is supported financially by the University of Indonesia, with DRPM PIT-9 2019 research grant scheme, ID number NKB-0039/UN2.R3.1/HKP05.00/2019

### References

[1] Versusarthritis accessed 14 July 2019 see https://www.versusarthritis.org/about-arthritis/conditions/osteoarthritis-of-the-knee/

[2] Orthoinfo accessed 14 July 2019 see https://orthoinfo.aaos.org/en/diseases--conditions/arthritis-of-the-knee/

[3] Webmd accessed 14 July 2019 see https://www.webmd.com/osteoarthritis/ostearthritis-of-the-knee-degenerative-arthritis-of-the-knee#1

[4] Adyalam T R, Rustam Z 2018 Classification of Osteoarthritis Disease Severity Using Adaboost Support Vector Machine *J. Phys. Conf. Ser.* **1108** (1) 012062

[5] Abdalla M E 2012 An Application on Multinomial Logistic Regression Model *Pakistan Journal of*
Statistics and Operation Research

[6] Cristianini N and Taylor J S 2000 An Introduction to Support Vector Machines and Other Kernel-based Learning Methods Cambridge University Press