



YÖNETİM BİLİŞİM SİSTEMLERİ DERGİSİ
<http://dergipark.ulakbim.gov.tr/ybs>

Yayın Geliş Tarihi: 24.11.2017
Yayına Kabul Tarihi: 31.11.2017
Online Yayın Tarihi: 20.12.2017

Cilt:3, Sayı:2, Yıl:2017, Sayfa:130-143
ISSN: 2148-3752

FORECASTING OF BOX OFFICE REVENUE USING MACHINE LEARNING ALGORITHMS

Özge Hüsniye NAMLI

Turk Alman University, Turkey

Tuncay ÖZCAN*

İstanbul University, Turkey

Abstract: Current developments such as new effects and 3D shootings increase the competition in the movie industry. Pre-production analyzes are becoming more important for the expensive and risky investments in the movie industry. At this point, the prediction of the box office revenue has become an important research issue. In this context, this study aims to present an approach using machine learning algorithms for box-office revenue prediction. Artificial neural networks and support vector machines algorithms as traditional artificial intelligence methods and random trees, random forests and C4.5 algorithms as decision tree algorithms are used. Later, a hybrid model is proposed using these algorithms and the bagging algorithm from the ensemble algorithm. Prediction models are evaluated with the percentage of correct classification, kappa statistics and ROC area. Numerical results show that Random forest-bagging and artificial neural networks-bagging hybrid methods have the best performance among all models.

Keywords: Box-office Revenue, Artificial Neural Networks, Support Vector Machines, Decision Trees, Machine Learning

MAKİNE ÖĞRENMESİ ALGORİTMALARI KULLANARAK GİŞE HASILATININ TAHMİNİ

Özet: Yeni efektler ve 3 boyutlu çekimler gibi güncel gelişmeler film endüstrisindeki rekabeti artırmaktadır. Film endüstrisindeki pahalı ve riskli yatırımlar için üretim öncesi analizler giderek önem kazanmaktadır. Bu noktada, gişe hasılatı tahmini önemli bir araştırma konusu olmuştur. Bu bağlamda, bu çalışma gişe hasılatı tahmini için makine öğrenmesi algoritmaları kullanarak bir yaklaşım sunmayı amaçlamaktadır. Geleneksel yapay zeka metodlarından yapay sinir ağları ve destek vektör makineleri algoritmaları, karar ağaçları algoritmalarından rastgele ağaç, rastgele orman ve C4.5 algoritmaları kullanılmıştır. Daha sonra, bu algoritmalar ile topluluk algoritmalarından torbalama algoritması kullanılarak melez bir model önerilmiştir. Tahmin modelleri doğru sınıflandırma yüzdesi, kappa istatistiği, ROC alanı ile değerlendirilmiştir. Sayısal sonuçlar,

rastgele orman-torbalama ve yapay sinir ağları-torbalama melez metotlarının tüm modeller arasında en iyi performansa sahip olduğunu göstermektedir.

Anahtar Kelimeler: Gişe hasılatı, Yapay sinir ağları, Destek vektör makineleri, Karar ağaçları, Makine öğrenmesi

*Contact Author: tuncay.ozcan@istanbul.edu.tr, İstanbul Üniversitesi, İstanbul, Türkiye

INTRODUCTION

The motion picture has completed their development over the years since 1800 and it has become a large industry which composes of directors, producers, screenwriters, actors, technical teams, distributors, auditoriums and audiences. Like many other fields, the movie industry is affected by developing technology and changing world. Movies require sizable investments for a lot of innovative process and creative work during the production. At the same time, Star crew, special effects, big decor, expensive costumes, etc., increase the cost of movies. In parallel with the costs, the budgets allocated for movies also are increased. At this point, the prediction of the box office revenue has become an important and risky problem. In this context, this study aims to present an approach using machine learning algorithms for box-office revenue prediction.

Many studies can be found for box-office revenue in the literature. These studies can be summarized as follows:

Sharda and Delen (2006) used neural networks to predict the box-office revenue of motion pictures before its theatrical release. They converted the forecasting problem into a classification problem consisting of nine categories instead of the point estimate of the box office revenues. The average percent hit rate was used as a performance measure to evaluate the results. The success of their neural network model is compared with the Multiple Logistic Regression (MLP), Discriminant Analysis (DA) and CART algorithms that previously proposed in the literature. The Neural Networks have shown much better results than others. Delen et al. (2007) created a web-based decision support system that uses the model in the study conducted by Sharda and Delen in 2006.

Zhang et al. (2009) developed a prediction model which uses multi-layer Back Propagation (MLBP) Neural Networks with multi-input and multi-output for forecasting box office revenue of a movie during the pre-production period. The input variables were defined using market survey and statistical methods were used to determine the weights of these input variables. A 6-fold cross validation was used to measure the performance of the proposed prediction model. When compared to the Multilayer Perceptron (MLP) method, the MLBP prediction model was found to give more satisfactory results. It is concluded that the use of MLBP prediction model is more reliable and effective for solving this problem.

Ghiassi et al. (2015) touched on that in the most of studies done to forecasting box office revenue in the film industry, existing models use the first week's data. This study set out with the aim of forecasting box-office revenue of movies before their theatrical release. Dynamic Artificial Neural Networks (DANN) method was applied by using the data at pre-production period. This study has compared their DANN model with SVM, ANN, Classification and Regression Trees (CART), RF and Moving Average in terms of average percent hit rate and found that the DANN is effective and increase the accuracy rate.

Hur et al. (2016) laid emphasis on forecasting box-office revenue for film distributors' decision-making process. They proposed a new box-office forecasting model that makes use of criticism/interpretation sensitivity and uses non-linear machine learning algorithms to increase the accuracy of prediction in their work. Multiple Linear Regression which is a linear model, CART, ANN, and SVR which are three different machine learning algorithms were used to reflect the nonlinear relationship between box-office revenues and forecasting algorithms. In addition to traditional forecasting algorithms, audience sentiments from criticism texts were used as input variable. In order to evaluate the proposed model in terms of prediction success, Root Mean Squared Error (RMSE) and mean absolute percentage error (MAPE) were utilized. The forecasting results have shown that machine learning algorithms give higher accuracy results when the available data is insufficient or long-term forecasts are made.

Castillo et al. (2016) focused on the problem of the number of books to be printed for delivery to bookstores. The problem of predicting total sales has dealt with the books to be printed in the right quantities yet before the books have distributed. In this study, the real data set containing all sales data for books published during seven years in Spain was used. The study was conducted in three stages. The first of these was the stage in which the data is examined with the aid of data visualization techniques. The other stage was feature selection and different techniques were used to determine which variables are more effective in sales at this stage. The final stage was the forecasting phase; the forecasting models were created using M5 model trees, RF, k-nearest neighbor algorithm, SVM, Linear Regression, MLP and extreme learning machine (ELM) algorithms for book sales at this stage. The performance of the applied algorithms was evaluated with mean absolute error (MAE), RMSE, Relative absolute error (RAE) and Root Relative Squared Error (RRSE). It has been observed that the forecasting algorithm giving the best results is M5A.

The purpose of this paper is to forecast the box office revenue of movies before its theatrical release. In parallel with this purpose, twelve different input variables are used. Box-office revenue as output variable is categorized in nine different groups, between the range of flops and blockbusters (Sharda & Delen, 2006). In the next parts of our study, detailed information about the dataset is given and the problem is defined. Then, 10-fold cross validation and 5-fold cross validation are used to divide the data set into training and testing sets. Artificial neural networks and support vector machines algorithms as traditional artificial intelligence methods and random trees, random forests and C4.5 algorithms as decision tree algorithms are applied as individual algorithms for forecasting box-office revenue. Later, a hybrid model is proposed using these algorithms and the bagging algorithm from the ensemble algorithm. Performance criteria such as the percentage of correct classification, kappa statistics, Precision, F-Measure, ROC Area and mean absolute error (MAE) are used to evaluate the prediction success of the algorithms. Numerical

results show that Random forest-bagging and artificial neural networks-bagging hybrid methods have the best performance among all models.

METHOD

In order to create the dataset used in this study, we procure the name, budget and box office revenue information of the movies from <https://boxofficeturkiye.com>. The remaining data is obtained according to the names of the movies from <http://www.imdb.com>. The data obtained from IMDb's website include information on year, MPAA ratings, release date, runtime, genre, director, actress/actor 1, actress/actor 2, language, country and metascore of the movies. Within the thirteen different variables of 3663 movies, we use year, MPAA, seasonality, runtime, genre, director, actress/actor 1, actress/actor 2, language, country, metascore and budget as input variables and box office revenue as output variable. In the light of all data on hand, we try to forecast the box office revenue of movies before its theatrical release

Detailed information about the variables used is given in the following table.

Table 1: Properties of variables

o	Variable	Definition of Variable	Type of variable	Variable Value
	Year	Release date of movie	Numerical	1927, ..., 2016
	MPAA	Motion picture rating	Categorical	G, PG, PG-13, R
	Seasonality	The rate of release month in terms of box-office revenue	Numerical	1, 2, 3, ..., 12
	Runtime	Runtime of movie	Numerical	20, 52, 63, ..., 213

	Genre	Genre of movie	Categorical	Action, Animation, Adventure, Biography, Comedy, Crime, Documentary, Drama, Family, Fantasy, History, Horror, Mysterious, Musical, Romantic, Science Fiction, Detective
	Director	Director's popularity	Categorical	0, 1, 2, ..., 6
	Actress/Actor 1	Star's popularity	Categorical	0, 1, 2, ..., 6
	Actress/Actor 2	Star's popularity	Categorical	0, 1, 2, ..., 6
	Language	Original language of movie	Categorical	English, German..
0	Country	The country where the movie was made	Categorical	England, USA, ...
1	MetaScore	The rate of movie in terms of popularity in social media	Categorical	1, 2, ..., 100
2	Budget	Budget spent for movie	Categorical	A, B, C, ..., I
3	Box-Office	Box-office revenue obtained from movie	Categorical	A, B, C, ..., I

Artificial Neural Network

Neural networks are the simulation model of the human nervous system, which consists of cells called neurons. These neurons are units of computation that take input from some other neurons, make computations on these inputs, and feed them into yet other neurons (Aggarwal, 2015). Artificial neural networks have the ability to "learn" mathematical relationships between input variables and output variables. This is achieved by "training" the network with a set of training data containing estimated variables and outcomes. Networks are programmed to set their internal weights based on mathematical relationships between inputs and outputs (Tu, 1996).

The ANN model is presented below (Erdal et al., 2013).

The output signal for the l th neuron in the n th layer is formulated as follows:

$$y_l^n(t) = \varphi \left[\sum_{j=1}^p w_{lj}^n(t) y_j^{n-1}(t) + \Psi_l^n \right] \quad (1)$$

Where $\varphi(\cdot)$ is the activation function, w_{lj}^n is the connection weight, t is the time index and $\Psi_l^n = w_{l_o}^n(t)$ is the weighted. The synaptic weight $w_{ji}^n(t)$ for an n -layer is calculated as follows:

$$w_{ji}^n(t+1) = w_{ji}^n(t) + \Delta w_{ji}^n(t) \quad (2)$$

Subject to $l \leq n \leq N$ and it can be revised as follows:

$$\Delta w_{ji}^n(t) = \eta \lambda_j^n(t) y_i^{n-1}(t) \quad (3)$$

Subject to $0 < \eta < 1$ where η denotes the learning rate, $\lambda_j^n(t) \equiv -\frac{\partial E_t}{\partial u_j^n}$ denotes the local error gradient. To advance the back-propagation algorithm, a momentum term α is added as in the following formula:

$$\Delta w_{ji}^n(t) = \eta \lambda_j^n(t) y_i^{n-1}(t) + \alpha \Delta w_{ji}^n(t-1) \quad (4)$$

Subject to $0 < \alpha < 1$

For the output layer, the local error gradient is given by:

$$\lambda_j^N(t) = [d_j(t) - y_j^N(t)] \varphi [u_j^N(t)] \equiv e_j(t) \varphi [u_j^N(t)] \quad (5)$$

Where $d_j(t)$ states the goal output signal and $\varphi(\cdot)$ states the activation function.

Support Vector Machine

The SVM is based on the principle of structural risk minimization. SVM can be analyzed theoretically using concepts from computational learning theory and achieve good performance in real world problems (Wu, Huang & Meng, 2008). Support vector machines are supervised learning models that select a small number of critical boundary instances called support vectors from each class, and form a linear discriminant function that separates them as much as possible (Witten, Frank & Hall, 2011).

The SVM algorithm works as follows (Chou et al., 2014).

The linear model in the space $f(x, w)$ can be formulated as follows:

$$f(x, \omega) = \sum_{j=1}^n w_j g_j(x) + b \quad (6)$$

$g_j(x)$ expresses a set of non-linear transformations from the input space, b is as a bias term, and w expresses the weight vector estimated by minimizing the regular risk function.

The loss function L_ε , which measures the quality of estimation, formulated below;

$$L_\varepsilon = L_\varepsilon(y, f(x, \omega)) = \begin{cases} 0, & \text{if } |y - f(x, \omega)| \leq \varepsilon \\ |y - f(x, \omega)|, & \text{otherwise} \end{cases} \quad (7)$$

SVM with ε -insensitive loss function, it calculates a linear regression function for the new high-dimensional feature space, and at the same time decreases the complexity of the model by reducing $\|\omega\|^2$ to the lowest. This function contains non-negative slack variables, ξ_i and ξ_i^* . Here $i = 1, \dots, n$ is used to describe training samples from the ε -insensitive zone.

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (8)$$

$$\text{subject to } \begin{cases} y_i - f(x_i, \omega) \leq \varepsilon + \xi_i^* \\ f(x_i, \omega) - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0, i = 1, \dots, n \end{cases}$$

The four core kernel functions used in SVM are as follows (Hsu et al., 2003);

Linear:

$$K(x_i, x_j) = x_i^T x_j \quad (9)$$

Polynomial:

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (10)$$

Radial Based Function:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (11)$$

Sigmoid:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \quad (12)$$

γ, r and d are the core parameters found in kernel functions.

Random Forest

Random Forest is an algorithm developed by Breiman (2001) and highly effective in estimation. Random Forest is an ensemble learning algorithm composed of many individual learners. It uses the bagging to create random sets for the decision tree setup. Although in the standard trees each node is branched using the best split among all variables, in the Random Forest each node is branched using the best among the subset of predictors randomly chosen at that node (Kalmegh, 2015).

Random Tree

Random Tree is a classification algorithm that generates a tree by taking randomly selected properties at a certain number of nodes at each node. There is no pruning and there is an option that allows predicting class possibilities based on the data set being retained (Akçetin & Çelik, 2014).

C4.5

The C4.5 decision tree building algorithm was developed by Quinlan (1993). C4.5, which is the evolution of ID3, uses gain ratio as splitting criteria and the splitting ceases when the number of instances to be split is below a certain threshold. When the dataset has numeric attributes, C4.5 can be used. It can also induce from a training set that incorporates missing values by using corrected gain ratio criteria (Rokach & Maimon, 2014).

Bagging Model

Bagging, which was proposed by Breiman (1996), is one of the ensemble machine learning algorithms. In the bagging algorithm, bootstrap samples of the training set are used to create the basic models and combined by plain voting for the classification task or averaging for the regression task (Cichosz, 2014).

The steps and description of the bagging model are as follows (Han & Kamber, 2006):

Steps of the Method:

1. for $i = 1$ to k do \rightarrow generate k models:
2. generate bootstrap sample, D_i , by sampling D with replacement;
3. use D_i to derivate a model, M_i ;
4. end for

To use the composite model on a tuple, X :

1. if classification then
2. let each of the k models classify X ;
3. return the majority vote;
4. end if
5. if prediction then
6. let each of the k models predict a value for X ;
7. return the average predicted value;
8. end

(D is a set of d training sample, k is model's number in the ensemble, M^* is a composite model)

Performance Criteria

The percentage of correct classification, kappa statistics, precision, F-Measure, ROC Area and MAE are used to evaluate and compare the performances of the algorithms.

The percentage of correct classification

The percentage of correct classification, which is one of the widely used performance criteria, shows the prediction success.

Table 2: The confusion matrix

	Positive (+)	Negative (-)
Positive (+)	True Positive (TP)	False Positive (FP)
Negative (-)	False Negative (FN)	True Negative (TN)

For calculate the percentage of correct classification, the following formula is used:

$$\begin{aligned} \text{The percentage of correct classification} \\ &= \frac{TP + TN}{TP + FP + FN + TN} \end{aligned} \quad (13)$$

Kappa Statistics

This statistic shows the correspondence between the actual values and the values of the classification model.

$$\begin{aligned} \kappa &= \frac{\sum O_{ij} - \sum E_{ij}}{n - \sum E_{ij}} \text{ for } i \\ &= j, O_{ij}: \text{ the observed values and } E_{ij}: \text{ expected values} \end{aligned} \quad (14)$$

Precision

Precision can be thought of as the probability that the detected structural change points are correct (Abou-Nasr et al., 2014).

$$\text{Precision} = \frac{TP}{TP + FP} \quad (15)$$

F-Measure

The F-measure is defined as a harmonic mean of precision and recall performance metrics (Sasaki, 2007).

$$F - measure = \frac{2Precision \times Recall}{Precision + Recall} \quad (16)$$

Area Under Receiver Operating Characteristic Curve

The area under the ROC curve is one of the commonly used performance metric. ROC curve shows the performance of a classifier. False positive rate is in the x-axis, and true positive rate is in the y-axis.

$$TP\ rate = \frac{TP}{TP + FN} \quad (17)$$

$$FP\ rate = \frac{FP}{TN + FP} \quad (18)$$

Mean Absolute Error

MAE is a measure of difference between two continuous variables. A better prediction success is obtained when the MAE is close to zero.

$$MAE = \frac{1}{N} \sum_{i=1}^N |x_{prep} - x_{obs}| \quad (19)$$

FINDINGS

The aforementioned algorithms are applied to movie dataset using the WEKA software. The results of the algorithms in terms of the performance criteria are shown below:

Table 3: Performance comparisons of the proposed models

		Cross-validation	Percentage of correct classification	Kappa Statistic	Precision	F-Measure	OC Area	Mean absolute error
Individual Classifiers	Random Tree	10 fold	88,3943	0,3307	0,885	0,885	0,733	0,0290
	Random Tree	5 fold	88,4762	0,3364	0,886	0,886	0,737	0,0291
	C4.5	10 fold	91,0705	0,3449	0,885	0,896	0,780	0,0316
	C4.5	5 fold	91,2343	0,3591	0,890	0,898	0,795	0,0315
	REPTree	10 fold	91,2616	0,3251	0,881	0,894	0,840	0,0313
	REPTree	5 fold	91,2889	0,3554	0,889	0,898	0,863	0,0308
	SVM	10 fold	91,5074	0,3496	0,885	0,897	,680	,1886
	SVM	5 fold	91,5074	,3472	0,886	0,897	,678	,1886
	ANN	10 fold	90,7428	,3842	0,892	0,899	,896	,0252

	ANN	5 fold	90,8520	,3855	0,893	0,900	,892	,0246
	Random Forest	10 fold	91,4801	,3682	0,890	0,900	,907	,0301
	Random Forest	5 fold	91,2889	,3587	0,889	0,898	,910	,0301
Ensemble Classifiers	Bagging (Random Forest)	10 fold	91,6166	,3660	0,8910	0,9000	,9140	,0303
	Bagging (Random Forest)	5 fold	91,3708	,3462	0,8880	0,8970	,9150	,0303
	Bagging (SVM)	10 fold	91,4528	,3422	0,8850	0,8960	,6980	,1871
	Bagging (SVM)	5 fold	91,2616	,3253	0,8830	0,8940	,6940	,1875
	Bagging (ANN)	10 fold	91,5074	,4094	0,8970	0,9040	,9140	,0262
	Bagging (ANN)	5 fold	91,2616	,3869	0,8930	0,9010	,9120	,0260

DISCUSSION AND CONCLUSIONS

In this study, the twelve different input variables with respect to the model output are used to forecast the box office revenue during the pre-production period. The data set is divided into training and testing sets using 10-fold cross validation and 5-fold cross validation. Artificial neural networks and support vector machines algorithms as traditional artificial intelligence methods and random trees, random forests and C4.5 algorithms as decision tree algorithms are applied as individual algorithms for forecasting box-office revenue. Later, a hybrid model is proposed using these algorithms and the bagging algorithm from the ensemble algorithm.

In the Table 3, the best values in each performance criteria are marked in bold. When the table containing the results of the algorithms in terms of the performance criteria is examined, it is seen that the bagging-random forest hybrid method gives the best value with 91.6166% in terms of percentage of correct classification. The best result for the kappa statistic, which shows the relation between real and predicted values, is achieved by ANN-bagging hybrid method. The ANN- bagging hybrid method also gives the best results for precision and F-measure. When the ROC area are examined, it is seen that the random forest-bagging hybrid method has a high accuracy. And finally, the ANN-bagging hybrid method gives the best results for MAE criteria.

When the results of the algorithms in terms of the performance criteria are examined, it is seen that the most successful algorithms for forecasting box-office revenue are random forest-bagging and ANN-bagging hybrid methods. The results

show that ensemble methods give better accuracy than individual methods for this problem.

In future studies, the prediction model can be created using the extreme learning machine algorithm, which was not previously applied in the box office revenue forecasting problem.

REFERENCES

Abou-Nasr, M., Lessmann, S., Stahlbock, R., & Weiss, G. M. (Eds.). (2014). *Real world data mining applications* (Vol. 17). Springer.

Aggarwal, C. C. (2015). *Data mining: the textbook*. Springer.

Akçetin, E., & Çelik, U. (2014). İstenmeyen Elektronik Posta (Spam) Tespitinde Karar Ağacı Algoritmalarının Performans Kıyaslaması. *Journal of Internet Applications & Management/İnternet Uygulamaları ve Yönetimi Dergisi*, 5(2).

Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123-140.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Castillo, P. A., Mora, A. M., Faris, H., Merelo, J. J., García-Sánchez, P., Fernández-Ares, A. J., ... García-Arenas, M. I. (2016). Applying Computational Intelligence Methods for Predicting the Sales of Newly Published Books in a Real Editorial Business Management Environment. *Knowledge-Based Systems*, 115, 133–151.

Chou, J. S., Tsai, C. F., Pham, A. D., & Lu, Y. H. (2014). Machine learning in concrete strength simulations: Multi-nation data analytics. *Construction and Building Materials*, 73, 771–780.

Cichosz, P. (2014). *Data Mining Algorithms: Explained Using R*. John Wiley & Sons.

Delen, D., Sharda, R., & Kumar, P. (2007). Movie forecast Guru: A Web-based DSS for Hollywood managers. *Decision Support Systems*, 43(4), 1151–1170.

Erdal, H. I., Karakurt, O., & Namli, E. (2013). High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform. *Engineering Applications of Artificial Intelligence*, 26(4), 1246–1254.

Ghiassi, M., Lio, D., & Moon, B. (2015). Pre-production forecasting of movie revenues with a dynamic artificial neural network. *Expert Systems with Applications*, 42(6), 3176–3193.

Han, J. & Kamber, M. (2006). *Data Mining: Concepts and Techniques*, San Francisco: Morgan Kaufmann Publishers.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification. Department of Computer Science, National Taiwan University.

Hur, M., Kang, P., & Cho, S. (2016). Box-office forecasting based on sentiments of movie reviews and Independent subspace method. *Information Sciences*, 372, 608–624.

Kalmegh, S. (2015). Analysis of WEKA Data Mining Algorithm REPTree, Simple Cart and RandomTree for Classification of Indian News. *International Journal of Innovative Science, Engineering & Technology*, 2(2), 438-446.

Quinlan, J. R. (1993). *C4. 5: Programs for machine learning*. California: Morgan Kauffmann Publishers.

Rokach, L., & Maimon, O. (2014). *Data mining with decision trees: theory and applications*. World scientific.

Sasaki, Y. (2007). The truth of the F-measure. *Teach Tutor mater*, 1(5).

Sharda, R., & Delen, D. (2006). Predicting box-office success of motion pictures with neural networks. *Expert Systems with Applications*.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology*, 49(11), 1225-1231.

Witten, I. H., Frank, E., & Hall, M. A., (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

Wu, T. K., Huang, S. C., & Meng, Y. R. (2008). Evaluation of ANN and SVM classifiers as predictors to the diagnosis of students with learning disabilities. *Expert Systems with Applications*, 34(3), 1846-1856.

Zhang, L., Luo, J., & Yang, S. (2009). Forecasting box office revenue of movies with BP neural network. *Expert Systems with Applications*, 36(3 PART 2), 6580–6587.