Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v14n1p266

Variable selection in gamma regression model using chaotic firefly algorithm with application in chemometrics

By Alkhateeb, Algamal

Published: 20 May 2021

Electronic Journal of Applied Statistical Analysis
Vol. 14, Issue 01, May 2021, 266-276
DOI: 10.1285/i20705948v14n1p266

Variable selection in gamma regression model using chaotic firefly algorithm with application in chemometrics

Ahmed Naziyah Alkhateeb[a] and Zakariya Yahya Algamal*[b]

[a] *Department of Operation Research and Intelligent Techniques, University of Mosul, Mosul, Iraq*
[b] *Department of Statistics and Informatics, College of Computers Sciences and Mathematics, University of Mosul, Iraq*

Published: 20 May 2021

Variable selection is a very helpful procedure for improving computational speed and prediction accuracy by identifying the most important variables that related to the response variable. Regression modeling has received much attention in several science fields. Firefly algorithm is one of the recently efficient proposed nature-inspired algorithms that can efficiently be employed for variable selection. In this work, chaotic firefly algorithm is proposed to perform variable selection for gamma regression model. A real data application related to the chemometrics is conducted to evaluate the performance of the proposed method in terms of prediction accuracy and variable selection criteria. Further, its performance is compared with other methods. The results proved the efficiency of our proposed methods and it outperforms other popular methods.

keywords: Variable selection, gamma regression model, firefly algorithm, chaotic map).

1 Introduction

Gamma regression is widely applied model for studying several real data problems, such as automobile insurance claims, healthcare economics, and medical science (De Jong and

*Corresponding author: zakariya.algamal@uomosul.edu.iq

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

Heller, 2008; Dunder et al., 2018; Malehi et al., 2015). "Specifically, gamma regression model is used when the response variable under the study is not distributed as normal distribution or the response variable is positively skewed. Consequently, the gamma regression assumes that the response variable has a gamma distribution" (Al-Abood and Young, 1986; Wasef Hattab, 2016).

In many real applications, recent developments in technologies have made the possibility to measure a large number of variables. In the regression modeling, the existence of huge number has a negative effect by overfitting the regression model. Therefore, identification of a small subset of important variables from a large number of variables set for accurate prediction is an important role for building predictive regression models (Algamal and Lee, 2015). When the number of variables increases, the traditional variable selection methods, such as stepwise selection, forward selection, and backward elimination computationally become an exhaustive search and require a long time for computing.

Recently, the naturally inspired algorithms, such as genetic algorithm, particle swarm optimization algorithm, firefly algorithm, and crow search algorithm, have a great attraction and proved their efficiency as variable selection methods (Sayed et al., 2019). This is because that the main target in variable selection is to minimize the number of selected variables while maintaining the maximum accuracy of prediction, and, therefore, they can be considered as optimization problems (Sindhu et al., 2017).

Several researchers have employed the naturally inspired algorithms for variable selection in regression models. Broadhurst et al. (1997) employed the genetic algorithm for variable selection in linear and partial least squares regression models, with application in chemometrics. Drezner et al. (1999) proposed to use tabu search algorithm in model selection in the linear regression model. On the other hand, a hybrid algorithm of genetic algorithm and simulated annealing was proposed as a subset selection method in linear regression model by Örkcü (2013). Brusco (2014) did a comparison of simulated annealing algorithms for variable selection in principal component analysis and discriminant analysis. Besides, the diferential evolution algorithm was used as a variable selection in linear regression model by Dunder et al. (2018). In generalized linear models, the natural inspired algorithms for variable selection are also used, such as, logistic regression model (Unler and Murat, 2010; Pacheco et al., 2009), Poisson regression model (Massaro and Bozdogan, 2015; Koç et al., 2018), and gamma regression model (Dunder et al., 2018).

The purpose of this paper is to propose chaotic firefly algorithm, which is a swarm intelligence technique, as an alternative variable selection method for use in gamma regression model. The proposed algorithm will efficiently help in identifying the most relevant variables in the count data regression model with a high prediction. The superiority of the proposed algorithm is proved though different simulation settings and a real data application in chemometrics.

The remainder of this paper organizes as follows. Section 2 covers the description of gamma regression models. The details of the firefly algorithm and the chaotic maps are illustrated in Sections 3 and 4, respectively. Section 5 is devoted to the real data application results. The conclusion is covered in Section 6.

2 Gamma regression model

Positively skewed data often arise in epidemiology, social, and economic studies. This type of data consists of nonnegative values. Gamma distribution is a well-known distribution that fits to such type of data. "Gamma regression model (GRM) is used to model the relationship between the positively skewed response variable and potentially regressors" (Uusipaikka, 2008).

Let y_i be the response variable and follows a gamma distribution with nonnegative shape parameter ν and nonnegative scale parameter γ, i.e. $y_i \sim Gamma(\nu, \gamma)$, then the probability density function is defined as:

$$f\left(y_i\right) = \frac{\gamma}{\Gamma\nu}(\gamma y_i)^{\nu-1}e^{-\gamma y_i}, \quad y_i \geq 0, \tag{1}$$

with $E(y) = \nu/\gamma = \theta$ and $var(y) = \nu/\gamma^2 = \theta^2/\nu$. Given that $\gamma = \nu/\theta$, Eq. (1) can re-parameterized as a function of the mean (θ) and the shape (ν) parameters and written depending on the exponential function as:

$$f\left(y_i\right) = EXP\left\{\frac{y_i(-1/\theta) - \log(-1/\theta)}{1/\nu} + c(y_i, \nu)\right\}, \tag{2}$$

where the canonical link function is $-1/\theta$, the dispersion parameter is $\phi = 1/\nu$ and $c(y_i, \nu) = \nu\log(\nu) + \nu\log(y_i) - \log(y_i) - \log(\Gamma(\nu))$.

Gamma regression model is usually modeled using the canonical link function (reciprocal), $\theta_i = -1/x_i^T\beta$ which is expressed as a linear combination of covariates$x_i = (x_{i1}, ..., x_{ip})^T$. The log link function, $\theta_i = \exp(x_i^T\beta)$, is alternatively used rather than the reciprocal link function because it ensures that $\theta_i > 0$.

The most common method of estimating the coefficients of GRM is to use the maximum likelihood method of Eq. (2). Given the assumption that the observations are independent and $\theta_i = -1/x_i^T\beta$, the log-likelihood function is given by:

$$\ell(\beta) = \sum_{i=1}^{n}\left\{\frac{y_i x_i^T\beta - \log(x_i^T\beta)}{1/\nu} + c(y_i, \nu)\right\}, \tag{3}$$

the ML estimator is then obtained by computing the first derivative of the Eq. (3) and setting it equal to zero, as:

$$\frac{\partial\ell(\beta)}{\partial\beta} = \frac{1}{\nu}\sum_{i=1}^{n}\left[y_i - \frac{1}{x_i^T\beta}\right]x_i = 0. \tag{4}$$

Unfortunately, the first derivative cannot be solved analytically because Eq. (4) is nonlinear in β. The iteratively weighted least squares (IWLS) algorithm or Fisher-scoring algorithm can be used to obtain the ML estimators of the gamma regression parameters. In each iteration, the parameters are updated by:

$$\beta^{(r+1)} = \beta^{(r)} + I^{-1}(\beta^{(r)})S(\beta^{(r)}), \tag{5}$$

where $S(\beta) = \partial\ell(\beta)/\partial\beta$ and $I^{-1}(\beta) = \left(-E\left(\partial^2\ell(\beta)/\partial\beta\partial\beta^T\right)\right)^{-1}$. The final step of the estimated coefficients is defined as:

$$\hat{\beta}_{GR} = (X^T\hat{W}X)^{-1}X^T\hat{W}\hat{u}, \tag{6}$$

where $\hat{W} = diag(\hat{\theta}_i^2)$ and $\hat{u}$ is a vector where ith element equals to $\hat{u}_i = \hat{\theta}_i + ((y_i - \hat{\theta}_i)/\hat{\theta}_i^2)$.

3 Firefly algorithm

In recent years, numerous nature-inspired algorithms have been proposed as powerful approaches to solve the continuous optimization problems. Minimizing the number of variables with maximizing the accuracy of prediction is an optimization problem (Sindhu et al., 2017).

Firefly optimization algorithm (FA) is one of the recently efficient proposed nature-inspired algorithms, which is firstly introduced by Yang Yang (2013). "The application of FA is an easy algorithm for solving the optimization problems compared with other algorithms. FA is inspired by the social behavior of fireflies through flashing lights. FA enables a swarm of fireflies with low light intensities to move towards the neighbor brighter fireflies possessing superior search abilities in solving optimization problems" (Al-Thanoon et al., 2019; Qasim et al., 2020; Kahya et al., 2019; Al-Thanoon et al., 2018; Algamal, 2019a; Al-Thanoon et al., 2020; Algamal, 2019b).

Three rules are held in FA (Yu et al., 2015). "The first rule is that all fireflies are unisex meaning that one firefly will be attracted to other fireflies regardless of their sex. The second rule is that the degree of the attractiveness of a firefly is proportion to its brightness, therefore for any two flashing fireflies, the less bright one will move towards the brighter one and the more brightness. If there is no brighter one than a particular firefly, it will move randomly. The third rule is that the brightness of a firefly is somehow related to the analytical form of the fitness function. For a maximization problem, the brightness of each firefly is proportional to the value of the cost function."

Let d represents the dimension of the object function that will optimized, n_f represents the number of fireflies, δ refers the light absorption coefficient, I_i is the light intensity, and r is the distance between any two firefly locations i (s_i) and j (s_j). This Cartesian distance can be defined as:

$$r(s_i, s_j) = \sqrt{\sum_{c=1}^{d}(s_{i,c} - s_{j,c})^2}. \tag{7}$$

Because I_i decreases when the distance from the source increases, the variations of I_i should be monotonically decreasing function. As a result, in most applications, the I_i can be approximated as:

$$I(r) = I_0\, e^{-\delta r^2}, \tag{8}$$

where I_0 is the original light intensity. Because the attractiveness of a firefly is proportional to the I_i, the attractiveness φ of a firefly is defined as:

$$\varphi(r) = \varphi_0\, e^{-\delta r^2}, \tag{9}$$

where φ_0 represents the attractiveness at $r = 0$. The movement of any firefly to the best position will be attracted to another firefly, which is more attractive firefly, by:

$$s_i^{(t+1)} = s_i^{(t)} + \varphi_0 \, e^{-\delta r_{i,j}^2} \left(s_j^{(t)} - s_i^{(t)} \right) + \alpha \, (k_1 - 0.5), \tag{10}$$

where α and k_1, respectively, is the randomization parameter and a generated random number from uniform distribution with [0, 1].

FA originally is proposed to solve continuous optimization problems. However, in variable selection, the optimization problem is discrete. A binary firefly algorithm (BFA) is proposed by Zhang, Gao (Zhang et al., 2016) to deal with the problem of variable selection where the position is binary. Because variable selection problem is to select a specific variable or not, thus the solution is expressed as a binary vector, where the value 1 indicates a variable to be selected and 0 otherwise. In BFA, the term

$$\varphi_0 \, e^{-\delta r_{i,j}^2} \left(s_j^{(t)} - s_i^{(t)} \right) + \alpha \, (k_1 - 0.5)$$

will transfer to probability vector by using the sigmoid (sigm) function as:

$$sigm = \frac{1}{1 + \exp\left[\varphi_0 \, e^{-\delta r_{i,j}^2} \left(s_j^{(t)} - s_i^{(t)} \right) + \alpha \, (k_1 - 0.5) \right]}. \tag{11}$$

Accordingly, the position of a firefly in Eq. (11) will be replaced as follow:

$$s_i^{(t+1)} = \begin{cases} 1 & if \ sigm \geq k_2 \\ 0 & otherwise, \end{cases} \tag{12}$$

where k_2 represents a random number generated from uniform distribution with [0, 1].

Table 1: A simple example

x_1	x_2	x_3	.....	$x_{\rho-1}$	x_ρ
1	0	0		1	0

4 The proposed chaotic firefly algorithm

Chaos theory describes erratic behavior in nonlinear systems and for this purpose, it uses chaotic maps. "Chaotic maps are visualized and can travel as particles in a limited range of nonlinear, dynamic, and nonlinear systems with no definite regularity-traveling path of these particles" (Sayed et al., 2018). Chaos strategy is applied to avoid being trapped in local optima and improve the quality of searching global optimum. Therefore, chaos has been employed in numerous optimization applications. Considering that the feature

selection problem is an optimization problem with searching range of [0, 1], chaos can be used to optimize this problem (Sayed et al., 2018).

In this paper, chaotic maps are considered to improve the performance of the binary firefly algorithm in terms of avoiding being trapped at the local optima and improving the convergence speed for variable selection in gamma regression model. Ten chaotic maps are used in this paper. The description of these maps is explained in Table 2.

Table 2: The description of the ten used maps

Name	Definition	Range
Chebyshev	$x_{k+1} = \cos(k \cos^{-1}(x_k))$	(-1,1)
Circle	$x_{k+1} = mod(x_k + 0.2 - \frac{0.5}{2\pi}\sin(2\pi x_k), 1)$	(0,1)
Guass/mouse	$x_{k+1} = \begin{cases} 1 & x_k = 0 \\ \frac{1}{mod(x_k,1)}, otherwise \end{cases}$	(0,1)
Iterative	$x_{k+1} = \sin\left(\frac{(0.7)\pi}{x_k}\right)$	(-1,1)
Logistic	$x_{k+1} = 4x_k(1 - x_k)$	(0,1)
Piecewise	$x_{k+1} = \begin{cases} \frac{x_k}{0.4} & 0 \le x_k < 0.4 \\ \frac{x_k - 0.4}{0.1} & 0.4 \le x_k < 0.5 \\ \frac{0.6 - x_k}{0.1} & 0.5 \le x_k < 0.6 \\ \frac{1 - x_k}{0.4} & 0.6 \le x_k < 1 \end{cases}$	(0,1)
Sine	$x_{k+1} = \sin(\pi x_k)$	(0,1)
Singer	$x_{k+1} = 1.07(7.86x_k - 23.31(x_k)^2 + 28.75(x_k)^3 - 13.302875(x_k)^4$	(0,1)
Sinusoidal	$x_{k+1} = 2.3x_k \sin(\pi x_k)$	(0,1)
Tent	$x_{k+1} = \begin{cases} \frac{x_k}{0.7} & x_k < 0.7 \\ \frac{10}{3}(1 - x_k) & x_k \ge 0.7 \end{cases}$	(0,1)

Consequently, our proposed algorithm setting is as follows:
Step 1: The number of fireflies is $n_f = 40$, $\varphi_0 = 1$, $\delta = 0.2$, $\alpha = 0.1$, and the maximum number of iterations is $t_{\max} = 500$.
Step 2: The positions of each firefly are randomly generated from uniform distribution with 0 and 1 for the original binary firefly algorithm. For the proposed chaotic maps, the maps which are described in Table 1 are used. The representation of the positions of a firefly is depicted in Figure 1.
Step 3: The fitness function is defined as:

$$fitness = \min\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right]. \tag{13}$$

Step 4: The positions of the fireflies are updated using Eq. (10).
Step 5: Steps 3 and 4 are repeated until a $t_{\max}$ is reached.

5 Real data application

To demonstrate the usefulness of the proposed method in real application, we present here a chemistry dataset with $(n, p) = (60, 4032)$, where n represents the number of chemical compounds. "While p denotes the number of molecular descriptors, which are treated as explanatory variables" (Algamal et al., 2015). The response of interest is the melting point. Quantitative structure-property relationship (QSPR) study has become a great deal of importance in chemometrics. The principle of QSPR is to model several physical activities over a collection of chemical compounds in terms of their structural properties (Algamal and Lee, 2017). Consequently, using of regression model is one of the most important tools for constructing the QSPR model.

The experimental values for the melting point (MP) in Kelvin (K) of 60 energetic carbocyclic nitroaromatic compounds were obtained from (Al-Fakih et al., 2018). The MP range of the compounds used is between 260.9 K and 489.1 K. The data were randomly divided into 42 compounds (70%) as a training dataset and 18 (30%) compounds as a test dataset. The molecular structures of the compounds were sketched using Chem3D software (CambridgeSoft Corporation, Cambridge, MA). The structures were optimized using the molecular mechanics (MM2) method implemented in the Chem3D software, and then using the molecular orbital package (MOPAC) module implemented in the same Chem3D software at the semi-empirical AM1 method, applying a minimum root mean square (RMS) gradient of 0.100 as a stopping criterion. DRAGON software (version 6.0) was used to generate 4885 molecular descriptors based on the optimized molecular structures.

To include consistent and useful descriptors, preprocessing steps were performed as follows. First, descriptors that had constant or zero values for all compounds were excluded (311 descriptors). Second, the remaining descriptors were further refned by removing those in which 70% of their values were zeros (255 descriptors). After that, descriptors with a relative standard deviation of less than 0.001 were removed (189 descriptors). In addition, the correlation of the remaining descriptors was examined to omit multicollinearity by removing those that were highly correlated ($r_{ij} \geq 0.90$) (98 descriptors). Finally, 4032 descriptors remained for constructing the QSPR model.

The performance of our proposed chaotic firefly algorithm maps is tested. Further, the performance of is compared with the original BFA. The performance is evaluated using the mean-squared error (MSE) of both train and test datasets, in addition, to the number of selected variables. The results are reported in Tables 2 and 3.

As it can be observed from Table 2, the binary firefly algorithm with different chaotic maps overtakes the BFA. In terms of the number of selected variables, most the chaotic maps selected few variables than the BFA. However, the Logistic map selected variables as same as the BFA. The best chaotic map is the Tent map, where it selected 9 variables out of 4032 compared to the other chaotic maps.

Moreover, it can be noticed that in the Tent map obtains the lowest MSE comparing with the other used chaotic maps. It can be seen that the MSE of the Tent map was about 25.509%, 23.428%, 26.208%, 31.981%, 31.002%, 18.116%, 5.916%, 2.478%, and 13.402% lower than that of Chebyshev, Circle , Guass, Iterative, Logistic, Piecewise, Sine, Singer, respectively. Additionally, comparing with the BFA, the Tent map was about 27.292% lower than BFA. It can be seen also that Singer map is the second best method. Regarding the test data (Table 3), it is obvious that the Tent map has the superiority of the results in term of MSE. The MSE of the Tent map was about 23.710%, 21.699%, 24.728%, 30.002%, 29.046%, 16.595%, 5.159%, 1.800%, and 12.100% lower than that of Chebyshev, Circle , Guass, Iterative, Logistic, Piecewise, Sine, Singer, respectively.

Table 3: The performance of the used methods for the train data

Method	# selected variables	MSE
BFA	24	4.221
Chebyshev	22	4.120
Circle	21	4.008
Guass	22	4.159
Iterative	23	4.512
Logistic	24	4.448
Piecewise	18	3.748
Sine	13	3.262
Singer	12	3.147
Sinusoidal	15	3.544
Tent	9	3.069

Table 4: The performance of the used methods for the test data

Method	MSE
BFA	4.465
Chebyshev	4.361
Circle	4.249
Guass/mouse	4.420
Iterative	4.753
Logistic	4.689
Piecewise	3.989
Sine	3.508
Singer	3.388
Sinusoidal	3.785
Tent	3.327

6 Conclusion

In this paper, the problem of selecting variables in gamma regression models is considered. A chaotic firefly algorithm with ten maps was proposed as a variable selection method. The results obtained from real data applications in chemometrics demonstrated the superiority of the chaotic firefly algorithm in terms of MSE for both the train and test data, and in addition, in terms of selected variables.

References

Al-Abood, A. and Young, D. (1986). Improved deviance goodness of fit statistics for a gamma regression model. *Communications in Statistics-Theory and Methods*, 15(6):1865–1874.

Al-Fakih, A., Algamal, Z., Lee, M., and Aziz, M. (2018). A penalized quantitative structure–property relationship study on melting point of energetic carbocyclic nitroaromatic compounds using adaptive bridge penalty. *SAR and QSAR in Environmental Research*, 29(5):339–353.

Al-Thanoon, N. A., Qasim, O. S., and Algamal, Z. Y. (2018). Tuning parameter estimation in scad-support vector machine using firefly algorithm with application in gene selection and cancer classification. *Computers in biology and medicine*, 103:262–268.

Al-Thanoon, N. A., Qasim, O. S., and Algamal, Z. Y. (2019). A new hybrid firefly algorithm and particle swarm optimization for tuning parameter estimation in penal-

ized support vector machine with application in chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 184:142–152.

Al-Thanoon, N. A., Qasim, O. S., and Algamal, Z. Y. (2020). Variable selection in gamma regression model using binary gray wolf optimization algorithm. In *Journal of Physics: Conference Series*, volume 1591, page 012036. IOP Publishing.

Algamal, Z. (2019a). Variable selection in count data regression model based on firefly algorithm. *Statistics, Optimization & Information Computing*, 7(2):520–529.

Algamal, Z. Y. (2019b). A particle swarm optimization method for variable selection in beta regression model. *Electronic Journal of Applied Statistical Analysis*, 12(2):508–519.

Algamal, Z. Y. and Lee, M. H. (2015). Penalized logistic regression with the adaptive lasso for gene selection in high-dimensional cancer classification. *Expert Systems with Applications*, 42(23):9326–9332.

Algamal, Z. Y. and Lee, M. H. (2017). A novel molecular descriptor selection method in qsar classification model based on weighted penalized logistic regression. *Journal of Chemometrics*, 31(10):e2915.

Algamal, Z. Y., Lee, M. H., Al-Fakih, A. M., and Aziz, M. (2015). High-dimensional qsar prediction of anticancer potency of imidazo [4, 5-b] pyridine derivatives using adjusted adaptive lasso. *Journal of Chemometrics*, 29(10):547–556.

Broadhurst, D., Goodacre, R., Jones, A., Rowland, J. J., and Kell, D. B. (1997). Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. *Analytica Chimica Acta*, 348(1-3):71–86.

Brusco, M. J. (2014). A comparison of simulated annealing algorithms for variable selection in principal component analysis and discriminant analysis. *Computational Statistics & Data Analysis*, 77:38–53.

De Jong, P. and Heller, G. Z. (2008). *Generalized linear models for insurance data*. Cambridge University Press.

Drezner, Z., Marcoulides, G. A., and Salhi, S. (1999). Tabu search model selection in multiple regression analysis. *Communications in Statistics-Simulation and Computation*, 28(2):349–367.

Dunder, E., Gumustekin, S., and Cengiz, M. A. (2018). Variable selection in gamma regression models via artificial bee colony algorithm. *Journal of Applied Statistics*, 45(1):8–16.

Kahya, M. A., Altamir, S. A., and Algamal, Z. Y. (2019). Improving firefly algorithm-based logistic regression for feature selection. *Journal of Interdisciplinary Mathematics*, 22(8):1577–1581.

Koç, H., Dünder, E., Gümüştekin, S., Koç, T., and Cengiz, M. A. (2018). Particle swarm optimization-based variable selection in poisson regression analysis via information complexity-type criteria. *Communications in Statistics-Theory and Methods*, 47(21):5298–5306.

Malehi, A. S., Pourmotahari, F., and Angali, K. A. (2015). Statistical models for the analysis of skewed healthcare cost data: a simulation study. *Health economics review*, 5(1):11.

Massaro, T. J. and Bozdogan, H. (2015). Variable subset selection via ga and information complexity in mixtures of poisson and negative binomial regression models. *arXiv preprint arXiv:1505.05229*.

Örkcü, H. H. (2013). Subset selection in multiple linear regression models: A hybrid of genetic and simulated annealing algorithms. *Applied Mathematics and Computation*, 219(23):11018–11028.

Pacheco, J., Casado, S., and Núñez, L. (2009). A variable selection method based on tabu search for logistic regression models. *European Journal of Operational Research*, 199(2):506–511.

Qasim, O. S., Al-Thanoon, N. A., and Algamal, Z. Y. (2020). Feature selection based on chaotic binary black hole algorithm for data classification. *Chemometrics and Intelligent Laboratory Systems*, 204:104104.

Sayed, G. I., Darwish, A., and Hassanien, A. E. (2018). A new chaotic whale optimization algorithm for features selection. *Journal of classification*, 35(2):300–344.

Sayed, G. I., Hassanien, A. E., and Azar, A. T. (2019). Feature selection via a novel chaotic crow search algorithm. *Neural Computing and Applications*, 31(1):171–188.

Sindhu, R., Ngadiran, R., Yacob, Y. M., Zahri, N. A. H., and Hariharan, M. (2017). Sine–cosine algorithm for feature selection with elitism strategy and new updating mechanism. *Neural Computing and Applications*, 28(10):2947–2958.

Unler, A. and Murat, A. (2010). A discrete particle swarm optimization method for feature selection in binary classification problems. *European Journal of Operational Research*, 206(3):528–539.

Uusipaikka, E. (2008). *Confidence intervals in generalized regression models*. Chapman and Hall/CRC.

Wasef Hattab, M. (2016). A derivation of prediction intervals for gamma regression. *Journal of Statistical Computation and Simulation*, 86(17):3512–3526.

Yang, X.-S. (2013). Multiobjective firefly algorithm for continuous optimization. *Engineering with Computers*, 29(2):175–184.

Yu, S., Zhu, S., Ma, Y., and Mao, D. (2015). Enhancing firefly algorithm using generalized opposition-based learning. *Computing*, 97(7):741–754.

Zhang, J., Gao, B., Chai, H., Ma, Z., and Yang, G. (2016). Identification of dna-binding proteins using multi-features fusion and binary firefly optimization algorithm. *BMC bioinformatics*, 17(1):323.