

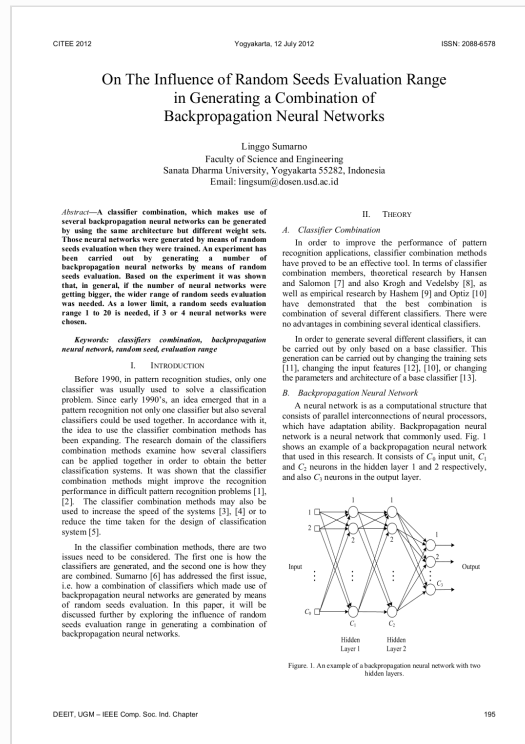


Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author: Sumarno Linggo
Assignment title: Periksa similarity
Submission title: On The Influence of Random Seeds Evaluation Range in Gen...
File name: n_Generating_a_Combination_of_Backpropagation_Neural_N...
File size: 562.89K
Page count: 6
Word count: 3,716
Character count: 18,617
Submission date: 07-Feb-2018 02:45PM (UTC+0700)
Submission ID: 912405082



On The Influence of Random Seeds Evaluation Range in Generating a Combination of Backpropagation Neural Network

by Linggo Sumarno

Submission date: 07-Feb-2018 02:45PM (UTC+0700)

Submission ID: 912405082

File name: n_Generating_a_Combination_of_Backpropagation_Neural_Network.pdf (562.89K)

Word count: 3716

Character count: 18617

On The Influence of Random Seeds Evaluation Range in Generating a Combination of Backpropagation Neural Networks

5
Linggo Sumarno

Faculty of Science and Engineering

Sanata Dharma University, Yogyakarta 55282, Indonesia

Email: lingsum@dosen.usd.ac.id

Abstract—A classifier combination, which makes use of several backpropagation neural networks can be generated by using the same architecture but different weight sets. Those neural networks were generated by means of random seeds evaluation when they were trained. An experiment has been carried out by generating a number of backpropagation neural networks by means of random seeds evaluation. Based on the experiment it was shown that, in general, if the number of neural networks were getting bigger, the wider range of random seeds evaluation was needed. As a lower limit, a random seeds evaluation range 1 to 20 is needed, if 3 or 4 neural networks were chosen.

Keywords: classifiers combination, backpropagation neural network, random seed, evaluation range

I. INTRODUCTION

Before 1990, in pattern recognition studies, only one classifier was usually used to solve a classification problem. Since early 1990's, an idea emerged that in a pattern recognition not only one classifier but also several classifiers could be used together. In accordance with it, the idea to use the classifier combination methods has been expanding. The research domain of the classifiers combination methods examine how several classifiers can be applied together in order to obtain the better classification systems. It was shown that the classifier combination methods might improve the recognition performance in difficult pattern recognition problems [1], [2]. The classifier combination methods may also be used to increase the speed of the systems [3], [4] or to reduce the time taken for the design of classification system [5].

In the classifier combination methods, there are two issues need to be considered. The first one is how the classifiers are generated, and the second one is how they are combined. Sumarno [6] has addressed the first issue, i.e. how a combination of classifiers which made use of backpropagation neural networks are generated by means of random seeds evaluation. In this paper, it will be discussed further by exploring the influence of random seeds evaluation range in generating a combination of backpropagation neural networks.

II. THEORY

A. Classifier Combination

In order to improve the performance of pattern recognition applications, classifier combination methods have proved to be an effective tool. In terms of classifier combination members, theoretical research by Hansen and Salomon [7] and also Krogh and Vedelsby [8], as well as empirical research by Hashem [9] and Optiz [10] have demonstrated that the best combination is combination of several different classifiers. There were no advantages in combining several identical classifiers.

In order to generate several different classifiers, it can be carried out by only base on a base classifier. This generation can be carried out by changing the training sets [11], changing the input features [12], [10], or changing the parameters and architecture of a base classifier [13].

B. Backpropagation Neural Network

A neural network is as a computational structure that consists of parallel interconnections of neural processors, which have adaptation ability. Backpropagation neural network is a neural network that commonly used. Fig. 1 shows an example of a backpropagation neural network that used in this research. It consists of C_0 input unit, C_1 and C_2 neurons in the hidden layer 1 and 2 respectively, and also C_3 neurons in the output layer.

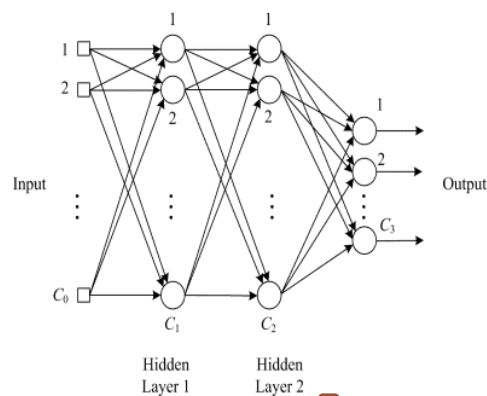


Figure 1. An example of a backpropagation neural network with two hidden layers.

7
1) *Training:* Neural network in Fig. 1 needs to be trained first, before it can be used in recognizing the input. Resilient propagation [14] is one of many training algorithms to train it.

2) *Initial weights:* One thing that carried out during the early step neural network training is assigning the initial weights of the neurons. The choice of the weights will influence the convergence rate or even the convergence failure of the neural network training [16].

- a) If the initial weights are set at the same values, the resulted error values will be constant over all training period. This situation will cause the neural network training trapped in the saturation that resist weights changing. Therefore, it can be judged that a convergence failure has been happened.
- b) If the initial weights are set at the different values (however they are inappropriate), it will cause a phenomenon called premature saturation [15]. This phenomenon refers to a situation where the resulted error values almost constant over a training period. This phenomenon cannot be judged as a local minimum, since error value will be reduced after that almost constant period over. This premature saturation phenomenon will slow down the convergence rate.

In order to avoid the two things above, in general practice, researchers used initial weights from random numbers that uniformly distributed in the small range [16].

3) *Random Numbers and Random Seeds:* One of computer property is deterministic property. Therefore, it cannot generate the real random numbers. Computer uses a pseudorandom generator, in order to mimic the real random number generator. By using this kind of generator, it can be generated a series of exact pseudorandom numbers, as long as the generator is initialized using the same initial number. This initial number called random seed.

When a process that make use a series of pseudorandom numbers is executed, it is possible to get an identical track record of the process. The neural network also makes use a series of pseudorandom numbers in the training process. Therefore, it is possible to get an identical track record of the training process, even though the training process is repeated again. On the other hand, by using a different series of pseudorandom numbers in the training process, it is possible to get a different track record of the training process. In the neural network training, a different track record means a different neural network, since it will has a different performance.

III. RESEARCH METHODOLOGY

A. Materials and Equipments

Materials in this research are isolated handwritten characters in binary format. These materials came from data acquisition sheets scanned at resolution of 300 dpi. The data were taken from 100 writers, from various levels of age (10 to 70 years) and sex. From 100 writers, each of them wrote 78 characters, which divided into three groups where each group consists of 'a' to 'z' characters. Therefore, there were 7,800 isolated character images. Equipments in this research was a set of computer based on processor Intel Core2Duo E7500 (2,93GHz) and 4GB RAM, that equipped with MATLAB software.

B. System Development

By using materials and equipments above, a system of handwritten character recognition has been developed (see Fig. 2). In that system, the input is an isolated character image in binary format, whereas the output is a character in the text format.

1) *Character Normalization:* Character normalization in Fig. 2 is carried out in order to correct problems of slant, size, shift, and stroke-width. In this research character normalization from Sumarno [17] was used. Fig. 3 shows some steps in character normalization.

In Fig. 3, the input is an isolated handwritten character in binary format, whereas the output is normalized handwritten character in binary format which has 64x64 pixels in size. Sumarno [17] suggested the following parameters.

- a) Slant correction was carried out by using evaluation of vertical projection histogram of handwritten character that had been undergone shearing operation by using shearing coefficients $\{-0.4, -0.35, \dots, 0.4\}$.
- b) Character scaling was set to 48x48 pixels.
- c) The template size was set to 64x64 pixels.
- d) Thinning operation used thinning algorithm from Zhang-Suen [18].
- e) Dilation operation used square structure-element 3x3.

9
2) *Feature Extraction:* Feature extraction is a process to extract features that exist in each of the character image. In this research feature extraction from Sumarno [17] was used. Fig. 4 shows some steps in feature extraction.

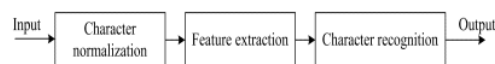


Figure 2. A character recognition system.

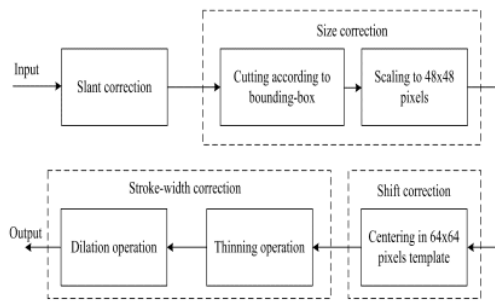


Figure 3. A character normalization steps.

In Fig. 4, the input is normalized character in binary format that has 64x64 pixels in size. The output is a set of values that represents the input image that has 64 elements. Sumarno [17] suggested the following parameters for the feature extraction steps.

- a) Low-pass filtering used 2D Gaussian filter 14x14 with standard deviation 10.
- b) Block partition used 8x8 pixels block size.

3) *Character Recognition*: Character recognition is a process to recognize a character. In order to recognize it, a recognition method based on a backpropagation neural network was used. Backpropagation neural networks in this research are described in detail as follow [17].

- a) Neural network with 2 hidden layers was chosen. It was chosen because after the evaluation of neural network with 1 and 2 hidden layers, the neural network with 2 hidden layers gave the highest recognition rate.
- b) Input layer has 64 neurons that correspond with the number of feature extraction elements.
- c) Output layer has 26 neurons that correspond with the number of alphabet characters 'a' to 'z'. Transfer function in this layer is unipolar sigmoid, that correspond with the network output i.e. in the range of 0 to 1.
- d) The hidden layers 1 and 2 have 64 and 312 neurons respectively. The number of neurons in each hidden layer were found from an evaluation procedure, where by using 64 and 312 neurons in hidden layer 1 and 2 respectively, it gave the highest recognition rate.
- e) Transfer functions in each hidden layer is bipolar sigmoid, that correspond with internal data processing in neural network which in the range -1 to 1.

Remarks

- a) Sigmoid function is a function that commonly used in a backpropagation neural network [20].
- b) Training of a backpropagation neural network can be more effective by using bipolar data processing in the range of -1 to 1 [19].



Figure 4. Feature extraction steps.

Patterns in training and testing

Patterns that used in training and testing the neural network were images of isolated handwritten character, which come from 100 persons that further processed into three pattern sets as follows [17].

a) Training set

This set used in training (in updating the neuron's weights). This set consists of 13,000 patterns as follows.

- i) There are 2,600 corrected patterns from group 1.
- ii) There are 5,200 corrected patterns from group 2. They come from corrected patterns from group 2 that rotated -5^0 and 5^0 .
- iii) There are 5,200 corrected patterns from group 3. They come from corrected patterns from group 3 that rotated -10^0 and 10^0 .

Remarks

- i) Every group consists of 2,600 patterns.
- ii) Corrected patterns are original patterns that have undergone slant, size, shift, and stroke width corrections.
- iii) It was assumed that the rotation in input patters is in the range of -10^0 to 10^0 .

b) Validation set

This set also used in training (in stopping the training process). They consist of 2,600 corrected patterns from group 2.

c) Test set

This set used in testing the trained neural network. They consist of 2,600 corrected patterns from group 3.

Training algorithm

The neural network was trained by using resilient backpropagation algorithm [14]. This algorithm is the fastest algorithm for pattern recognition [21]. Stopping criterion in training was made use of validation, in order to avoid under-training or over-training.

Pseudorandom and random seed

Neural network needs random initial weights in order to start the training. The computer generates these initial weights. However, since the computer cannot generate the real random numbers, therefore the pseudorandom numbers were used as initial weights. In this research, that pseudorandom numbers have the following properties.

- a) Distribution : uniform [16].
- b) Range : -1 to 1 (since bipolar sigmoid function has limit numbers -1 and 1).
- c) Repeatability : 2^{1492} (built-in in the MATLAB software).

Initial weights in each layer of neural network should be different, in order to remove correlation between layers. Therefore, random seed values that used in generating pseudorandom numbers in each layer should also be different. In this research, random seed values that used in hidden layers 1, 2 and output layer were i (i are integer numbers), $i+1$ and $i+2$ respectively [6].

IV. TESTING AND DISCUSSIONS

Experiments below were carried out by using a number of backpropagation neural networks, which have the same architecture and also the same training sets. The difference in each experiment only in terms of random seeds evaluation range. In the first experiment, there were 5 neural networks that trained using random seeds evaluation range 1 to 5. Table I shows the results of the first experiment. Based on the Table I(B), if we choose top 3, 4, and 5 neural networks, they will give average recognition rate 87.68, 86.56, and 86.25 respectively.

In the second experiment, Table I was expanded by training 10 neural networks using random seeds evaluation range 1 to 10. Table II shows the results of the second experiment. Based on the Table II(B), if top 3, 4, and 5 neural networks are chosen, they will give average recognition rate 87.08, 86.98, and 86.88 respectively.

In the third experiment, Table II was expanded by training 15 neural networks using random seeds evaluation range 1 to 15. Table III shows the results of the third experiment. Based on the Table III(B), if top 3, 4, and 5 neural networks are chosen, they will give average recognition rate 87.26, 87.12, and 87.03 respectively.

TABLE I. (A) TRAINING RESULTS OF 5 NEURAL NETWORKS THAT TRAINED USING RANDOM SEEDS EVALUATION RANGE 1 TO 5; (B) SORTING OF (A) BASED ON RECOGNITION RATE IN DESCENDING MANNER

(A)		
Neural network number	Random seed	Recognition rate (%)
1	1	87.39
2	2	87.15
3	3	85.00
4	4	86.69
5	5	85.00

(B)		
Neural network number	Recognition rate (%)	Random seed
1	87.39	1
2	87.15	2
3	86.69	4
4	85.00	3
5	85.00	5

In the fourth to sixth experiments there were 15, 20, 25, and 30 neural networks that trained using different random seeds evaluation ranges. They were trained using random seeds evaluation ranges of 1 to 20, 1 to 25, and 1 to 30 respectively. The results of top 3, 4, and 5 chosen neural networks in terms of average recognition rate are shown in the Table IV. The results of random seed evaluation ranges 1 to 5 and 1 to 10 are shown also in the Table IV.

TABLE II. (A) TRAINING RESULTS OF 10 NEURAL NETWORKS THAT TRAINED USING RANDOM SEEDS EVALUATION RANGE 1 TO 10; (B) SORTING OF (A) BASED ON RECOGNITION RATE IN DESCENDING MANNER.

(A)		
Neural network number	Random seed	Recognition rate (%)
1	1	87.39
2	2	87.15
3	3	85.00
4	4	86.69
5	5	85.00
6	6	86.38
7	7	85.62
8	8	86.69
9	9	86.31
10	10	86.46

(B)		
Neural network number	Recognition rate (%)	Random seed
1	87.39	1
2	87.15	2
3	86.69	4
4	86.69	8
5	86.46	10
6	86.38	6
7	86.31	9
8	85.62	7
9	85.00	3
10	85.00	5

TABLE III. (A) TRAINING RESULTS OF 15 NEURAL NETWORKS THAT TRAINED USING RANDOM SEEDS EVALUATION RANGE 1 TO 15.

(A)		
Neural network number	Random seed	Recognition rate (%)
1	1	87.39
2	2	87.15
3	3	85.00
4	4	86.69
5	5	85.00
6	6	86.38
7	7	85.62
8	8	86.69
9	9	86.31
10	10	86.46
11	11	82.08
12	12	85.62
13	13	85.62
14	14	86.00
15	15	87.23

TABLE III. (CONTINUED) (B) SORTING OF (A) BASED ON RECOGNITION RATE IN DESCENDING MANNER.

(B)		
Neural network number	Recognition rate (%)	Random seed
1	87.39	1
2	87.23	15
3	87.15	2
4	86.69	4
5	86.69	8
6	86.46	10
7	86.38	6
8	86.31	9
9	86.00	14
10	85.62	7
11	85.62	12
12	85.62	13
13	85.00	3
14	85.00	5
15	82.08	11

Based on the Table IV, it can be seen that if 3 or 4 neural networks are chosen, at least a random seeds evaluation range 1 to 20 is needed, whereas if 5 to 7 neural networks are chosen, at least a random seeds evaluation range 1 to 25 is needed. These cases are happened because certain random seed values that give high recognition rates are randomly distributed. However, there are limits in its distribution range. For example, for 3 and 4 neural networks, the limit is at 20 random seeds (for random seeds evaluation range 1 to 20), whereas for 5 to 7 neural networks, the limit is at 25 random seeds (for random seeds evaluation range 1 to 25).

Based on the Table IV, in general, it can be said that if more neural networks are chosen, the wider range of random seeds evaluation is needed. As a lower limit, a random seeds evaluation range 1 to 20 is needed, if 3 or 4 neural networks are chosen.

V. CONCLUSION

Based on the above discussions, there are two conclusions as follow.

- A new study about the influence of random seeds evaluation range in choosing a number of backpropagation neural networks as classifier combination members has been carried out. In this case, these neural networks have the same architecture but different weight sets.
- In general, it can be said that if more neural networks are chosen, the wider range of random seeds evaluation is needed. As a lower limit, a random seeds evaluation range 1 to 20 is needed, if 3 or 4 neural networks are chosen.

TABLE IV. AVERAGE RECOGNITION RATE OF CHOSEN NEURAL NETWORKS (IN %)

Number of chosen neural networks	Random seeds evaluation range					
	1 to 5	1 to 10	1 to 15	1 to 20	1 to 25	1 to 30
3	87.08	87.08	87.26	87.28	87.28	87.28
4	86.56	86.98	87.12	87.25	87.25	87.25
5	86.25	86.88	87.03	87.22	87.23	87.23
6	-	86.79	86.94	87.13	87.21	87.21
7	-	86.72	86.86	87.07	87.16	87.16
8	-	86.59	86.79	87.01	87.11	87.13
9	-	86.41	86.70	86.95	87.06	87.08
10	-	86.27	86.59	86.89	87.02	87.05

Remarks: random seed evaluation range 1 to 5, 1 to 10, 1 to 15, 1 to 20, 1 to 25, and 1 to 30 are correspond with the number of evaluated neural networks 5, 10, 15, 20, 25, and 30 respectively.

REFERENCES

- [1] J. Kittler and F. Rolli, editors, Proceeding of the First International Workshop on Multiple Classifier Systems, Cagliari, Italy, 2000.
- [2] J. Kittler and F. Rolli, editors, Proceeding of the Second International Workshop on Multiple Classifier Systems, Cambridge, UK, 2001.
- [3] Y. Chim, A. Kasim, and Y. Ibrahim, "Dual classifier system for handprinted alphanumeric character recognition", Pattern Analysis and Applications, vol. 1, 1998, pp. 155-162.
- [4] L. Lam and C. Y. Suen, "Structural classification and relaxation matching of totally unconstrained handwritten zip-code numbers", Pattern Recognition, vol. 21, no. 1, 1998, pp. 19-31.
- [5] S. Gunter, Multiple Classifier Systems in Offline Cursive Handwriting Recognition, PhD Thesis, Institute for Computer Science and Applied Mathematics, University of Bern, German, 2004, unpublished.
- [6] L. Sumamo, "Generating a combination of backpropagation neural networks by means of random seeds evaluation", Proceeding of 12th International Conference on QIR (Quality in Research), Bali, Indonesia, 2011, pp. 282-287.
- [7] L. Hansen, and O. Salomon, "Neural network ensembles". IEEE Transactions on Pattern Recognition and Machine Analysis, vol. 12, no. 10, 1990, pp. 993-1001.
- [8] A. Krogh, and J. Vedelsby, "Neural networks ensembles, cross validations, and active learning", Advances in Neural Information Processing Systems, G. Tesauro et.al. (editor), MIT Press, vol. 7, 1995, pp. 231-238.
- [9] S. Hashem, "Optimal linear combinations of neural networks", Neural Networks, vol. 10, no. 4, 1997, pp. 599-614.
- [10] D. W. Optiz, "Feature selection for ensembles", Proceeding of 16th International Conference on Artificial Intelligence, 1999, pp. 379-384.
- [11] L. Breiman, "Bagging predictors", Machine Learning, vol. 24, no. 2, 1996, pp. 123 - 140.
- [12] T.K. Ho, "The random subspace method for constructing decision forests", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, 1998, pp.832 - 844.
- [13] D. Partridge, and Y. B. Yates, "Engineering multiversion neural-net systems", Neural Computation, vol. 8, no. 4, 1996, pp. 869-893.
- [14] M. Riedmiller, and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm", Proceedings of the IEEE International Conference on Neural Networks, 1993, pp. 586-591.
- [15] Y. Lee, S. Oh, and M.Kim, "The effect of initial weights on premature saturation in backpropagation learning", International

- Joint Conference on Neural Networks, Seattle, Washington, USA, vol. 1, 1991, pp.765-770.
- [16] S. Haykin, Neural Networks : A Comprehensive Foundation, New Jersey: Prentice-Hall International Inc., 1994.
- [17] L. Sumarno, "On the performance of blurring and block averaging feature extraction based on 2D Gaussian filter", Proceeding of 5th International Conference on Information and Communication Technology and Systems (ICTS), Surabaya, Indonesia, 2009, pp. 261-266.
- [18] T.Y. Zhang, T.Y. and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns", Communication of The ACM, vol. 27, no. 3, 1984, pp. 236-239.
- [19] I. Suhardi, Evaluation of Artificial Neural Network for Handwritten Character Recognition Handprinted Style (Evaluasi Jaringan Syaraf Tiruan untuk Pengenalan Karakter Tulisan Tangan Jenis Cetak), Master Thesis, Electrical Engineering, Gadjah Mada University, Yogyakarta, Indonesia, 2003, unpublished.
- [20] L. Fausett, Fundamentals of Neural Networks, New Jersey: Prentice Hall International Inc., 1994.
- [21] The Mathworks Inc., Neural Network Toolbox: For Use With MATLAB, Version 5, Massachussets: The Mathworks Inc., 2005.

On The Influence of Random Seeds Evaluation Range in Generating a Combination of Backpropagation Neural Network

ORIGINALITY REPORT

7%

SIMILARITY INDEX

4%

INTERNET SOURCES

2%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Universitat Politècnica de València

Student Paper

2%

2

repository.gunadarma.ac.id

Internet Source

1%

3

archive.org

Internet Source

1%

4

Gunter, S.. "Off-line cursive handwriting recognition using multiple classifier systems- on the influence of vocabulary, ensemble, and training set size", Optics and Lasers in Engineering, 200503/05

Publication

<1%

5

Agustinus Rudatyo Himamunanto, Anastasia Rita Widiarti. "Javanese character image segmentation of document image of Hamong Tani", 2013 Digital Heritage International Congress (DigitalHeritage), 2013

Publication

<1%

6	Submitted to Universitas Pancasila Student Paper	<1 %
7	Xia-ting Feng, M. Seto, K. Katsuyama. "Neural dynamic modelling on earthquake magnitude series", Geophysical Journal International, 1997 Publication	<1 %
8	Studies in Computational Intelligence, 2008. Publication	<1 %
9	www.hindawi.com Internet Source	<1 %
10	A. F. R. Rahman, M. C. Fairhurst. "Multiple classifier decision combination strategies for character recognition: A review", International Journal on Document Analysis and Recognition, 2003 Publication	<1 %
11	dspace.lboro.ac.uk Internet Source	<1 %
12	pure.tue.nl Internet Source	<1 %
13	www.ijert.org Internet Source	<1 %

Exclude quotes On

Exclude bibliography On

Exclude matches

< 5 words