Chapter 7

Artificial Neural Network Based Fault Analysis

7.1 Introduction

This chapter contains a detailed description of ANN based fault analysis. This chapter starts with the topology of ANN model, followed by the methodology used in fault analysis of drag system and constructs of the ANN model. Finally, the results and discussions of cause to symptom and symptom to fault model are described.

7.2 Topology of Developed ANN Based Fault Analysis

The ANN is a very powerful artificial intelligence tool to identify the faulty pattern and classify the fault by pattern recognition [56], [276], [277]. ANN model consists of a large number of interconnected artificial neurons whose problem-solving capabilities involve learning from multiple input parameters and their associated output parameters [34], [55], [278]. Due to the complex nonlinear relationship between cause to symptom and symptom to fault of the drag system, ANN model is used for analysis. It is analysed using two ANN models. The first model is based on cause and symptom, and the second one is based on symptom and fault relationship. The cause to symptom ANN model is used to identify the sequence of responsible root causes to avoid the occurrence of fault when the symptom exceeds the threshold limit. The symptom to fault ANN model is used to predict the possible occurrence of faults in the drag system to avoid catastrophic failure whenever the symptom exceeds the threshold limit. Therefore, the combination of both models can be used in real-time application to prepare the effective CBM policy to avoid the occurrence of fault as well as failure of the drag system.

The ANN model consists of mainly three attributes: architecture or topology, transfer or activation function, and learning or training algorithm [279]. The architecture of the ANN model involves arranging the number of the hidden layers, number of neurons in each hidden layer and their connection, and it is used in the MLP and trained by backpropagation algorithm [280]–[282]. The activation function is used to solve the complex and nonlinear problem that allows to reach its best performance by controlling the output of each neuron in the hidden layer and output layer [33], [279]. A learning algorithm is a systematic step by step procedure through which the connection weights among neurons are adjusted to minimize the difference between the predicted and actual output of the model and controlled by the learning parameter such as learning rate and momentum [243], [283], [284].

The knowledge that the network acquires during learning is implicitly encoded in its numeric weights and bias values. Each layer of the neural network acts as a bridge between the input and output parameters. A bias is an additional constant parameter in the neural network, which is used to adjust the output along with the weighted sum of the inputs to the neuron. The output of the MLP neural network is given in Eq: (7.1) [283]

$$Y_{jk} = f_k\left(\sum_{i=1}^{N_{k-1}} W_{ijk} Y_{i(k-1)} + b_{jk}\right) \tag{7.1}$$

where Y_{jk} and b_{jk} are the neuron of j^{th} output from k^{th} layer and bias weight for neuron j in layer k, respectively. The model-fitting parameters, W_{ijk}, are the link weights that were selected randomly at the beginning of the network training process, and f_k is the nonlinear activation transfer functions.

7.3 Methodology of ANN Based Fault Analysis

For the development of ANN models, the collected data of the drag system was assigned randomly to the training, testing, and holdout subsets. The training data subset of the ANN model is used to build the ANN model architecture and adjust the weights and biases. The testing data subset is used for calculating the errors, and the holdout data subset is used to validate the ANN model [11]. Detail description of the methodology to optimize the ANN model for fault prediction and to adopt it in the preventive maintenance policy of dragline is depicted in Figure 7.1.

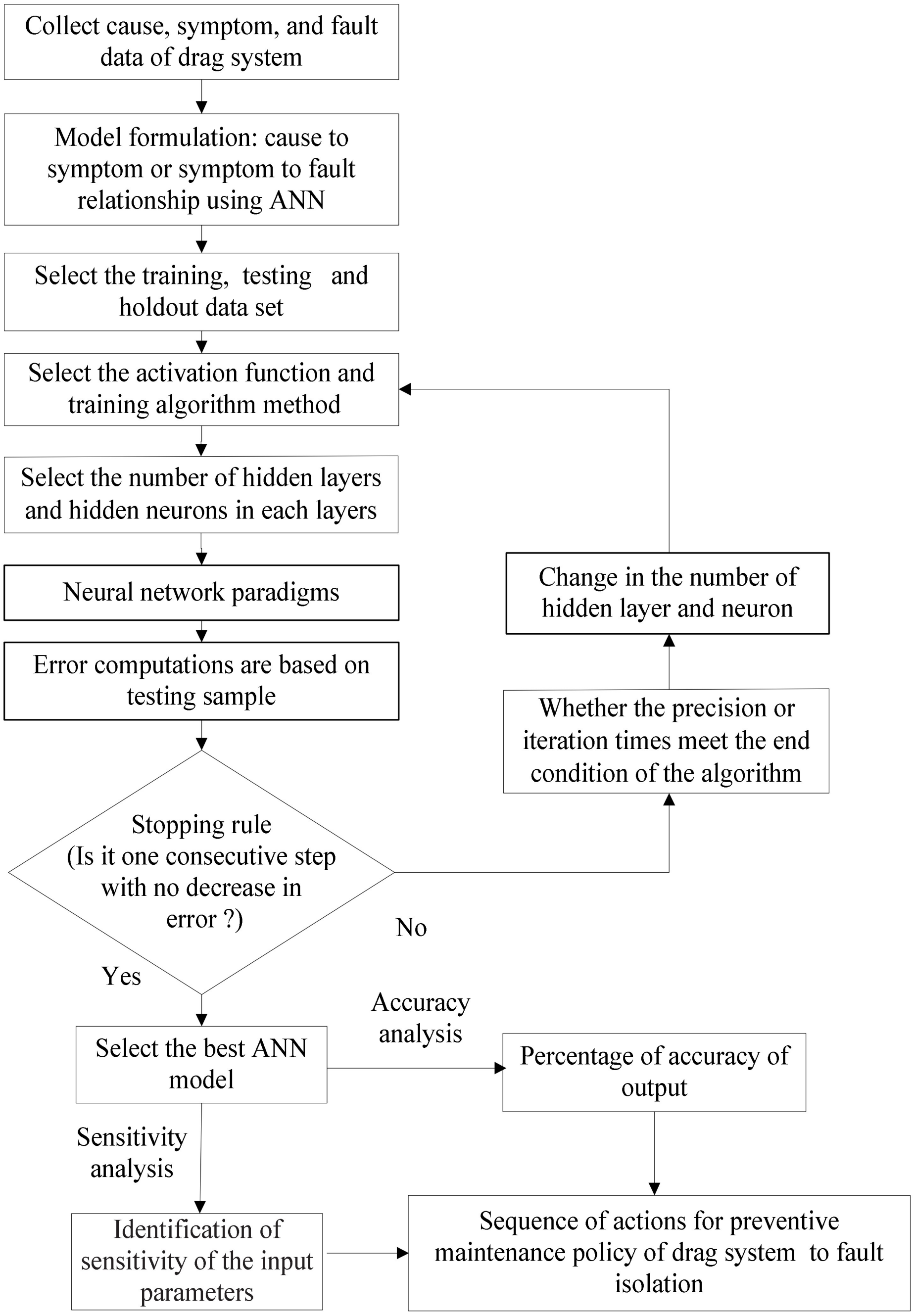

Figure 7.1 Applied algorithms to create an optimized ANN and maintenance policy

In the ANN model, initially the number of hidden layers and number of hidden neurons in each layer were selected based on the rule of thumb [285]. When testing a range of hidden neurons, it is important to keep all other parameters constant (e.g. learning rate, momentum). For a fixed number of training cycles, several neural network configurations are developed, varying only the number of neurons in the hidden layer. The optimal configuration is the one achieving minimum error during the limited training session.

The ANN model is used to perform the classification of target value using softmax activation for the output layer and cross-entropy error function to predict the accuracy of the model. Cross-entropy error is directly proportional to the difference between the predicted and actual output values of the ANN model. The cross-entropy loss increases as the predicted probability diverge from the actual level. Cross-entropy error (E) in the backpropagation algorithm to reduce stagnation is calculated using Eq. (7.2) [286].

$$E = -\sum_{J=1}^{n} (Y_{pj} \log Y_{oj} - (1 - Y_{pj}) \log(1 - Y_{pj})) \tag{7.2}$$

where Y_{pj} is the target output and Y_{oi} actual output of the model, n is the number of class. The cross-entropy error function in the network tries to minimize the error during training. The ANN model finds the accuracy of each output of the model and makes the sequence of the importance of the input parameters for each output and the sequence of the maintenance actions in the maintenance policy.

7.4 Modelling of Collected Drag System Data

The available collected data of the drag system is pre-processed, when one or more than one symptom crosses the threshold limit. When the symptom crosses the threshold limit, the corresponding occurrences of fault and source of fault (i.e., cause) are recorded. Some of the causes and faults occur without giving any symptom is also considered in this study, which is named as an unidentified symptom (US). There were 14 such data samples (out of 452) that did not give any symptom, but the fault had occurred. These are categorized as unidentified symptoms. These faults occur mainly due to the effect of other subsystems of a dragline (i.e., sudden fall of the bucket, boom that affects the drag system), and these faults are categorized into an unidentified fault (UF). When the cause of the occurrence of a fault is unknown, they are categorised into the unidentified cause (UC). Hence three more variables for UC, US and UF are considered in the fault analysis based on the ANN model. In order to understand the process of data collection, two examples of drag system fault have been cited.

☐ When EHMR was 46:00, the sudden increase in temperature due to overheating of the drive control system was warned through the alarm, due to the degraded nature of the fault; hence dragline continued to operate. The maintenance was done during the potential to failure (P-F) interval to minimize downtime and maintenance costs. The data retrieval from the maintenance record, as mentioned in Table 7.1, is as follows:

- Categorisation of symptom, TF = 1, remaining symptoms are within the threshold limit, hence CF = VF = VUS = LL= Sp = US = 0
- Categorisation of cause data, OH = 1 which was identified during maintenance and remaining causes were not observed during maintenance, hence BLB = IC = LC = OL = DC = UC = 0

- Occurrence of fault DCP = 1, therefore remaining faults did not occur in this case and the fault was identified during maintenance. Hence BBF = SBP = ID = UF = 0.

❑ When EHMR was 326:45, the vibration and unwanted sound were observed and during maintenance, the cause, bearing and loose bolt was identified. The replacement of bearing was done during P-F interval to minimise the downtime. Hence symptom VUS = 1, and remaining symptoms, which were within the threshold limit, were assigned '0' value. The cause BLB = 1 and remaining causes were not observed during maintenance. Fault BBF = 1, and remaining faults were not observed during maintenance.

Table 7.1 Cause, symptom, and fault data of the drag system and corresponding EHMR

Sl. No.	EHMR	Cause							Symptom							Fault				
		BLB	IC	OL	LC	OH	DC	UC	CF	VF	VUS	LL	TF	Sp	US	BBF	DCP	SBP	ID	UF
1.	46:00	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
2.	326:45	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
3.	745:30	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4	938:15	0	0	1	0	1	1	0	0	0	0	0	1	1	0	0	1	0	0	0
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
451	7071:15	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1
452	7267:30	0	1	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0

There were 452 instances when the symptoms crossed the threshold limit. Out of 452 cases, 199 symptoms were converted into faults. For the purpose of ANN modeling, these 452 data has been transformed to categorical data, as presented in Table 7.1, along with their corresponding EHMR values. There were 16 failures in the drag system as realized from these 199 faults.

7.5 Implementation of the ANN Model and Parameters Estimation

The MLP module of ANN models was used to build the neural network model and to test the accuracy and sensitivity of each parameter using SPSS v.19 statistical package [258]. The prediction accuracy of the ANN model provides reasonable confidence to diagnostic results, which is presented in Eq. (7.3) [287].

$$Prediction\ Accuracy = \frac{CP_1 + CP_0}{CP_1 + ICP_1 + CP_0 + ICP_0} \tag{7.3}$$

Here, CP_1 and CP_0 represent the correct prediction of the output parameters, i.e., observed value is '1' predicted as '1' and observed value is '0' predicted as '0' respectively. The ICP_1 and ICP_0 representing the incorrect prediction of the output parameter, i.e. observed value is '1' predicted as '0', and observed value is '0' is predicted as '1' respectively.

7.5.1 Cause to symptom model development

The cause to system ANN model is developed using cause nodes as input parameter and symptom node as the output parameter. The collected data was analyzed in the SPSS software to make the relationship between cause and symptom. In SPSS, the MLP procedure uses random number generation during random assignment of partitions, random subsampling for initialization of synaptic weights, random subsampling for automatic architecture selection, and the simulated annealing algorithm is used in weight initialization and automatic architecture selection. To reproduce the same randomized results in the future, the same initialization value for the random number generator is used before each run of the MLP procedure. Simulated annealing process is used to optimize the parameters in a model. This process is very useful for situations where there are a lot of local minima such that algorithms like gradient descent would be stuck at. Simulated annealing can be computationally heavy if it is tasked with many iterations, but it is capable of finding a global maximum and not stuck at local minima [288]. The collected 452 data are assigned in three groups for model development and validation: 273 (60.4%) data were used as training dataset, 82 (18.1%) as testing data and 97 (21.5%) as holdout data. The random division of data in training, testing, and holdout dataset measures are tested for statistical similarities using analysis of variance (ANOVA) [289]–[291]. In one-way ANOVA, the F-statistic is calculated to show the variance between the means of three groups of dataset are not significantly different. The F-test is performed at a significant level $\alpha = 0.05$.

- Null hypothesis (H_0): There is no difference in means of the dataset subgroups.
- Alternative hypothesis (H_1): At least any two means are not all equal.

As stated earlier, the statistical similarities of the three datasets are verified using ANOVA. The F-test statistics is used to compare the three dataset in order to prove that all the three dataset have similar properties before they were analysed. The F-critical value for the given dataset is 3.016. The mean, standard deviation, and p-value of the individual parameters of cause and symptom parameters for training, testing and holdout dataset are presented in Table 7.2.

Table 7.2 Statistical summary of the random division of the cause to symptom data set

Parameter	Training dataset (N=273)		Testing dataset (N=82)		Holdout dataset (N=97)		F-value	p-value
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation		
BLB	0.209	0.407	0.219	0.416	0.206	0.406	0.028	0.973
IC	0.586	0.493	0.610	0.491	0.639	0.483	0.432	0.649
OL	0.190	0.393	0.207	0.407	0.175	0.382	0.147	0.863
LC	0.366	0.483	0.341	0.477	0.309	0.465	0.524	0.593
OH	0.476	0.500	0.488	0.503	0.525	0.502	0.351	0.705
DC	0.029	0.169	0.036	0.188	0.031	0.174	0.055	0.946
UC	0.139	0.347	0.122	0.329	0.164	0.373	0.352	0.703
CF	0.374	0.485	0.293	0.458	0.371	0.486	0.932	0.392
VF	0.106	0.309	0.159	0.367	0.072	0.260	1.734	0.178
VUS	0.161	0.368	0.219	0.416	0.206	0.407	0.975	0.378
LL	0.212	0.409	0.171	0.379	0.268	0.445	1.276	0.280
TF	0.509	0.501	0.427	0.498	0.546	0.500	1.336	0.264
Sp	0.319	0.467	0.317	0.468	0.361	0.483	0.311	0.733
US	0.154	0.361	0.183	0.389	0.134	0.342	0.407	0.666

N* = Number of data sample

The null hypothesis is not rejected as the calculated F-value is below the F-critical value. The p-value of all parameters from Table 7.2 also reveals that the null hypothesis (H_0) is not rejected. This result provides that the generation data subsets i.e., training data, testing data and holding data used in the ANN model is statistically similar.

The activation function of the models in the hidden layer is hyperbolic tangent and that of the output layer is softmax. The performance of the ANN model is controlled by the learning parameters, including learning rate and momentum. The number of neurons is optimized during the training process by the trial and error method. The relationship between the average cross-entropy error and variation on the number of neurons is shown in Figure 7.2. Therefore 11 neurons are chosen as the average cross-entropy error is minimum. The schematic representation of the cause to symptom ANN model is shown in Figure 7.3.

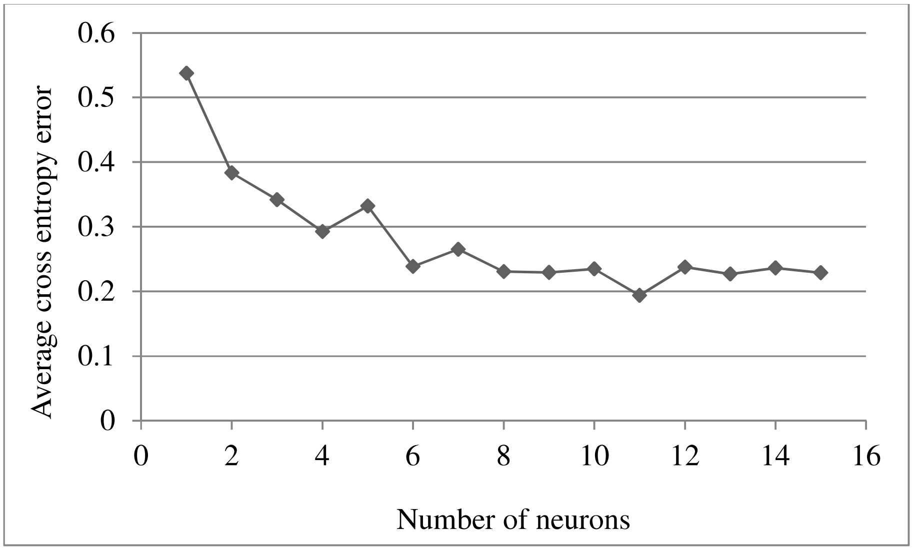

Figure 7.2 Relation between average cross-entropy error and number of neurons

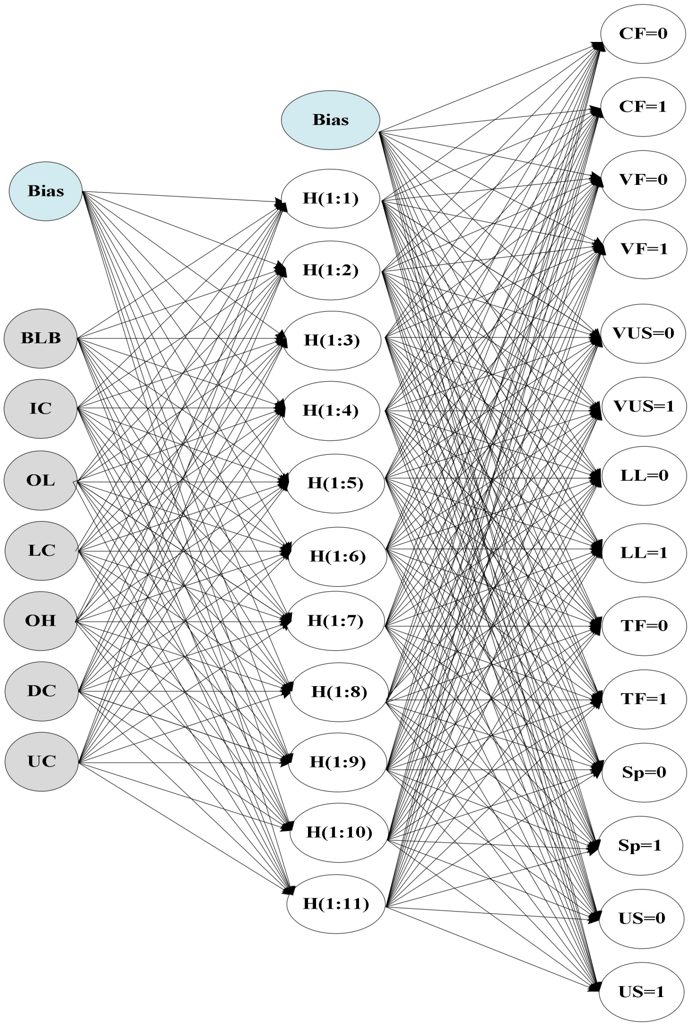

Figure 7.3. Multilayer perceptron cause to symptom ANN model

A learning rate or the training-rate coefficient adjusts the size of the average weight change and momentum, and it can also be included with a portion of the last weight change to accelerate the training convergence and improve the training precision [279]. In these models, architecture correctly identifies the high accuracy, having the learning rate coefficient (η) 0.4 and momentum (α) 0.9. These models are used to batch training to optimize the gradient descent method to minimize errors. The stopping rule used in these ANN models having one consecutive step with no decrease in error and error computations in the testing data subset.

7.5.2 Symptom to fault model development

The cause to system ANN model is developed using symptom nodes as input parameter and fault node as the output parameter. The collected data is fed in SPSS software using to make the relationship between symptom and fault. Out of the collected 452 data, 266 (58.8%) data are used for training, 94 (20.8%) for testing, and 92 (20.4%) as holdout data. The mean, standard deviation, and p-value of the individual parameters of cause and symptom parameter for training, testing and holdout dataset are presented in Table 7.3.

Table 7.3 Statistical summary of the random division of the symptom to fault data set

Parameter	Training dataset (N=266)		Testing dataset (N=94)		Holdout dataset (N=92)		F-value	p-value
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation		
CF	0.379	0.486	0.330	0.473	0.327	0.471	0.636	0.529
VF	0.101	0.303	0.127	0.336	0.109	0.313	0.244	0.783
VUS	0.177	0.382	0.191	0.395	0.185	0.390	0.055	0.946
LL	0.222	0.416	0.202	0.404	0.196	0.400	0.176	0.839
TF	0.504	0.501	0.500	0.502	0.500	0.503	0.003	0.997
Sp	0.323	0.469	0.340	0.476	0.326	0.471	0.046	0.955
US	0.154	0.362	0.159	0.368	0.152	0.361	0.011	0.989
BBF	0.116	0.321	0.096	0.296	0.109	0.313	0.154	0.857
DCP	0.105	0.307	0.117	0.323	0.089	0.283	0.229	0.795
SBP	0.172	0.378	0.096	0.296	0.120	0.326	1.987	0.138
ID	0.086	0.282	0.106	0.309	0.087	0.283	0.176	0.839
UF	0.157	0.365	0.085	0.281	0.130	0.339	1.577	0.208

It can be concluded from Table 2 that the F-value of all parameters is below the F-critical value and it reveals that the null hypothesis (H_0) is not rejected. Therefore the

ANOVA results reveals the generation of statistically similar data subsets (i.e., training data, testing data and holding data) those are used in symptom to fault ANN model.

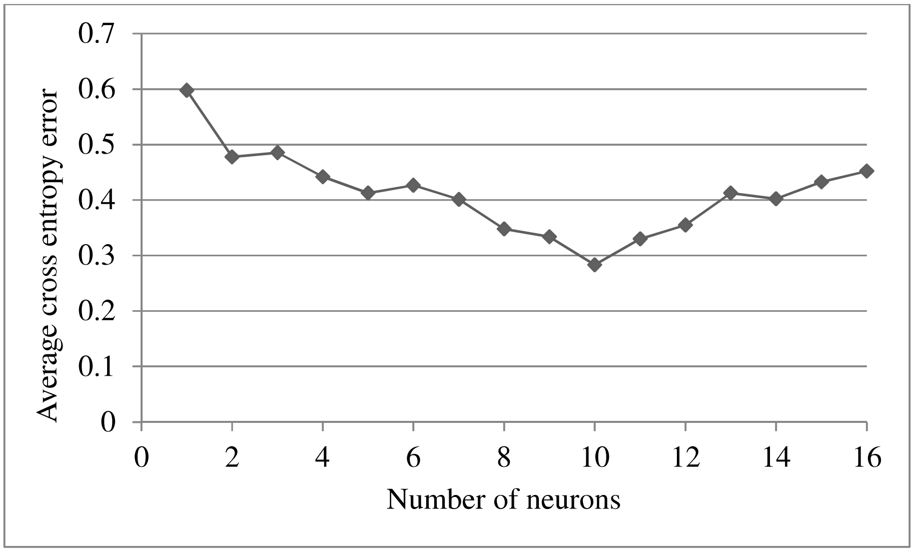

Figure 7.4 Relation between average cross-entropy error and number of neurons

The activation function of the hidden layer is hyperbolic tangent and that of the output layer is softmax. In these models, architecture correctly identifies the high accuracy, having the learning rate coefficient (ɳ) 0.4 and momentum (α) 0.9. These models are used to batch training to optimize the gradient descent method to minimize errors. The stopping rule used in these ANN models having one consecutive step with no decrease in error and error computations in the testing data subset. The number of neurons is optimized during the training process by the trial and error method. The relationship between the average cross-entropy error and variation on the number of neurons is shown in Figure 7.4. The schematic representation of the symptom to fault ANN model is shown in Figure 7.5.

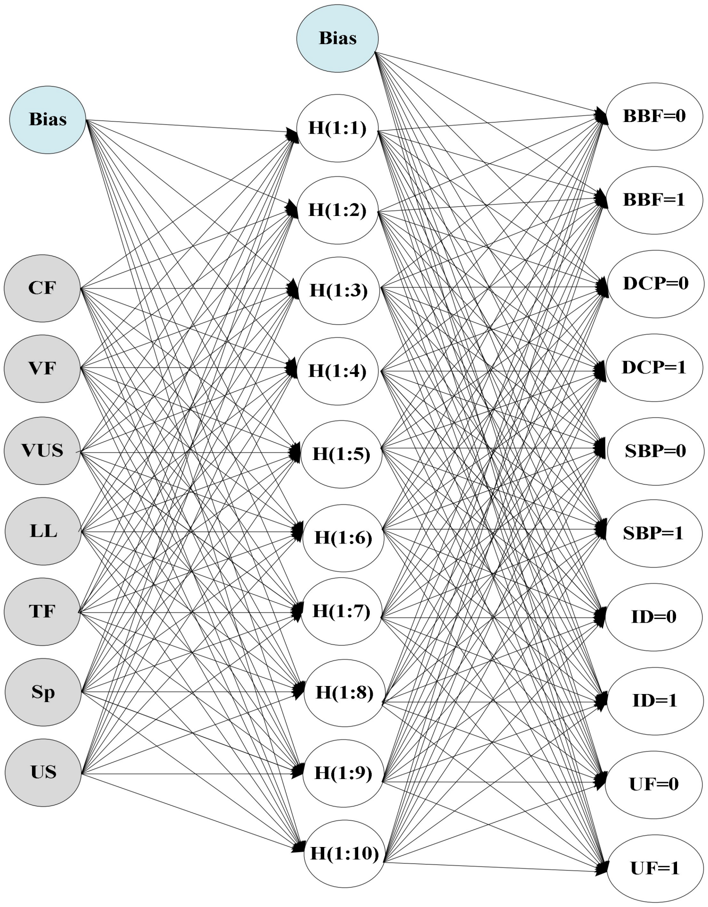

Figure 7.5 Multilayer perceptron symptom to fault ANN model

7.6 Results and Discussions

7.6.1 Cause to symptom ANN model and its prediction accuracy

The predicted pseudo-probabilities of cause to symptom ANN model is depicted in Figure 7.6. Each fault parameter has two categories when it crossed the threshold limit, it is referred as ‘1’ and when it is within the threshold limit, it is referred as ‘0’. The threshold values of causes and symptoms are presented in section 5.3. The X-axis represents the output parameter observed response categories (i.e., 0 or 1) and the Y-axis corresponds to predicted pseudo-probability [292]. The objective of showing the output in a box plot is to display the predicted output distribution of ANN model to easily compare with actual output within a limited space to visualize the results.

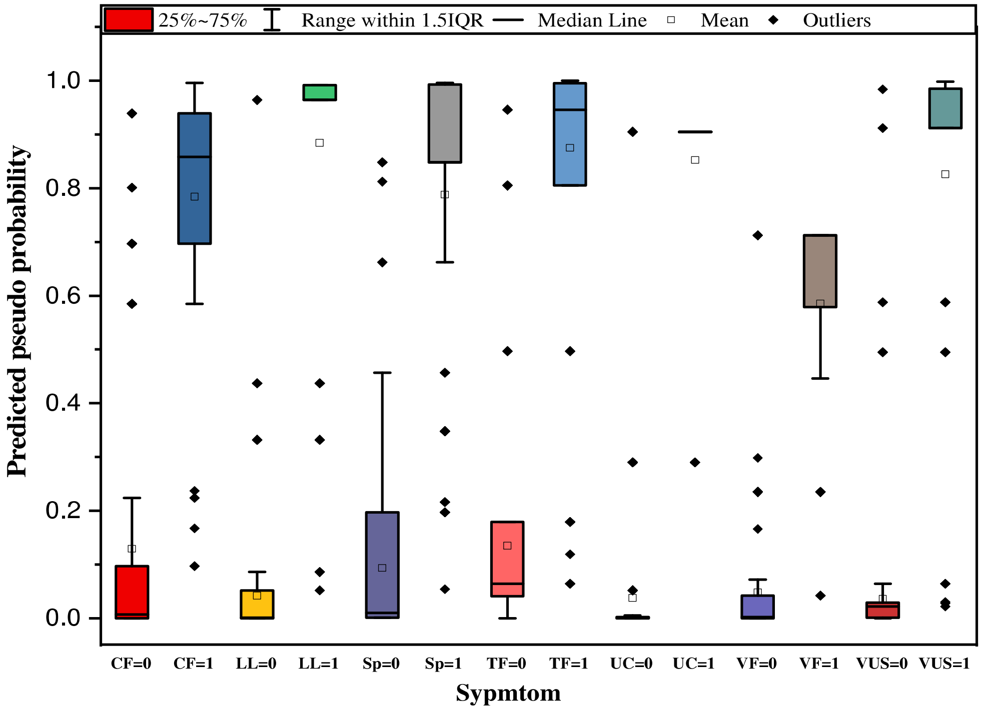

Figure 7.6 Predicted pseudo-probability of the output of cause to symptom model

Box plot displays the batches of data and spacing between the different parts of the box, indicating the degree of dispersion (spread) and skewness in the output of the ANN model. It also shows the outliers. The interquartile range (IQR) is a measure of variability, and based on variability the dataset can be divided into quartiles. Quartiles divide the rank-ordered data set into four parts such as interquartile range, midhinge, mid-range, and trimean. Five values from a set of data are conventionally used: the extremes, the upper and lower hinges (quartile) and the median. Hence the viewer can determine the confidence interval around the medians in a relative manner [293]. In the observed categories of ‘0’ in the box plot represents predicted pseudo-probability below the 0.5 marks is correct prediction and above the 0.5 mark represents the incorrect prediction. In the output categories of ‘1’ represents predicted pseudo-probability above the 0.5 marks is correct prediction and below the 0.5 mark represent the incorrect prediction. The number of correct and incorrect predictions of the output of three data subset, such as training, testing and holdout data subset, of cause to symptom model is presented in Table 7.4. In the ANN model, the percentage accuracy of prediction is expected to provide confidence to model results.

Table 7.4. Classification of the predicted output of cause to symptom analysis and its prediction accuracy

Symptom	Observed	Training data set			Testing data set			Holdout data set			Overall prediction accuracy
		Predicted		Prediction accuracy	Predicted		Prediction accuracy	Predicted		Prediction accuracy	
		0	1		0	1		0	1		
CF	0	144	27	0.842	50	8	0.862	52	9	0.852	0.897
	1	4	98	0.961	0	24	1.000	1	35	0.972	
VF	0	234	10	0.959	68	1	0.986	87	3	0.967	0.959
	1	8	21	0.724	1	12	0.923	1	6	0.857	
VUS	0	220	9	0.961	62	2	0.969	77	0	1.000	0.979
	1	3	41	0.932	2	16	0.889	2	18	0.900	
LL	0	212	3	0.986	68	0	1.000	71	0	1.000	0.979
	1	10	48	0.828	1	13	0.929	2	24	0.923	
TF	0	126	8	0.940	45	2	0.940	39	5	0.886	0.907
	1	16	123	0.885	2	33	0.943	4	49	0.925	
Sp	0	178	8	0.957	55	1	0.982	61	1	0.984	0.897
	1	17	70	0.805	6	20	0.769	9	26	0.743	
US	0	224	7	0.970	67	0	1.000	83	1	0.988	0.979
	1	4	38	0.905	1	14	0.933	1	12	0.923	

The cause and symptom relationship can be used to identify the sequence of causes, when any symptom occurs in the drag system. In the real-time condition, when the symptom vibration and unwanted sound exceeds the threshold limit and the causes are unknown, then the sequence to identify the causes based on the sensitivity analysis can be made. It is observed that the most effective parameter for symptom vibration and unwanted is the drum and coupling having weightage 34.4%, hence it should be inspected first (Table 7.5). When the drum and coupling are not in good conditions, the gearbox, drum and soft alignment, or rope are not aligned properly to the groove, and accordingly necessary action of maintenance should be taken. Otherwise, the maintenance engineer will move to the second most effective parameter that is the bearing jam and loose bolt, having weightage 24.2% for inspection.

Table 7.5. Rank-wise importance of cause and symptom of the drag system

Cause	Symptom weightage(rank)						
	CF (rank)	VF (rank)	VUS (rank)	LL (rank)	TF (rank)	Sp (rank)	US (rank)
BLB	0.081 (7)	0.032 (7)	0.242 (2)	0.145 (4)	0.091 (5)	0.063 (7)	0.112 (4)
IC	0.175 (3)	0.196 (2)	0.128 (3)	0.175 (2)	0.190 (3)	0.163 (3)	0.187 (2)
OL	0.187 (1)	0.092 (5)	0.123 (5)	0.056 (6)	0.198 (2)	0.156 (4)	0.027 (7)
LC	0.117 (5)	0.181 (3)	0.017 (7)	0.039 (7)	0.074 (7)	0.199 (1)	0.174 (3)
OH	0.185 (2)	0.111 (4)	0.021 (6)	0.331 (1)	0.216 (1)	0.140 (5)	0.082 (6)
DC	0.106 (6)	0.079 (6)	0.344 (1)	0.083 (5)	0.080 (6)	0.097 (6)	0.112 (5)
UC	0.151 (4)	0.306 (1)	0.125 (4)	0.171 (3)	0.150 (4)	0.182 (2)	0.305 (1)

If the bearing and bolt are not in good conditions, the necessary action of maintenance such as tightening of bolts, or friction of bearing can be done. After that, the engineer should move to the third most effective parameter that is the improper control (weightage 17.5%), and it should be inspected. Immediate action needs to be taken for improper control due to operator mistakes, including improper speed control, braking, and pushing wrong switches. If it is found to be in proper order, then go to the fourth most effective parameter that is the unidentified causes (weightage 12.5%) and the unidentified cause occurs randomly and it is difficult to identify, because they would not provide any symptom/warning to the drag system. Some of these failures are failure of boom or broken rope that affects the drag system as well as the whole system of the dragline. When it does not occur, then go to the fifth most effective cause parameter that is the overloading (weightage 11.7%) hence it should be inspected. After that go to the least effective parameter, that is the loose connection (weightage 1.7%). It is because of the electrical connectivity of the system and it occurs mainly due to the broken wire, short-circuits or open-circuit. Similarly, the sequence of maintenance policy of the drag system when any other symptom occurs to identify the cause can be made. In Table 7.5, the relationship between cause to symptom and the corresponding rank and the contribution is presented. The maintenance policy can be made to reduce the occurrence of a major fault in the drag system.

Similarly, when the symptom current feedback crosses the threshold limit, it is observed that most effective cause is the overloading having weightage 18.7%, hence it should be inspected for the overloading. The second most effective parameter is overheating (weightage 18.5%). It occurs mainly due to improper lubrication on bearing hence the lubrication of the bearing can be inspected, and if it is found to be in good condition, then go to the third most effective cause and so on depending on the weightage of cause for every symptom in the drag system.

The average time to inspect one parameter takes about 5 to 20 min as reported by the dragline maintenance staff. In the proper sequence, most of the responsible causes can be identified after one or two steps, so it can reduce the downtime and it can improve the production and safety of the dragline. It will also help to take preventive maintenance action before the fault occurrence.

7.6.2 Symptom to fault ANN model and its predicted accuracy

The predicted pseudo-probabilities of the symptom to fault ANN model is shown in Figure 7.7. The correct and incorrect prediction of the output of three data subset, such as training, testing and holdout data subset, of symptom to fault model is presented in Table 7.6. In the ANN model, the percentage accuracy of prediction will give confidence to the maintenance engineer when making the preventive maintenance action on the drag system of the dragline.

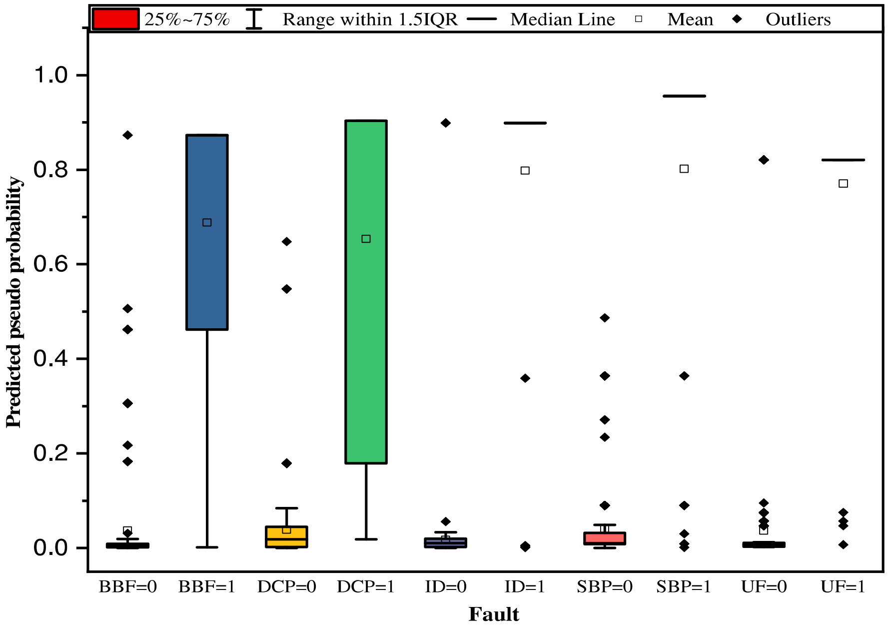

Figure 7.7 Predicted pseudo-probability of the output of symptom to fault model

The symptom to fault relationship is used to identify the sequence that of the symptoms, when any fault in the drag system is identified. In the case study, for instance, when the fault bolt or bearing failure are identified, then a sequence to identify the symptom (Table 7.7) can be made, and it can then be linked to the cause (Table 7.7).

Table 7.6 Classification of the predicted output of symptom to fault analysis and its prediction accuracy

Faults	Observed	Training data set			Testing data set			Holdout data set			Overall prediction accuracy
		Predicted		Prediction accuracy	Predicted		Prediction accuracy	Predicted		Prediction accuracy	
		0	1		0	1		0	1		
BBF	0	234	1	0.996	84	1	0.988	80	2	0.976	0.935
	1	13	18	0.581	2	7	0.778	4	6	0.600	
DCP	0	238	0	1.000	83	0	1.000	84	0	1.000	0.989
	1	11	17	0.607	3	8	0.727	1	7	0.875	
SBP	0	220	0	1.000	84	1	0.988	81	0	1.000	0.978
	1	10	36	0.783	0	9	1.000	2	9	0.818	
ID	0	241	2	0.992	83	1	0.988	84	0	1.000	0.989
	1	2	21	0.913	1	9	0.900	1	7	0.875	
UF	0	214	10	0.955	86	0	1.000	78	2	0.975	0.978
	1	4	38	0.905	0	8	1.000	0	12	1.000	

Table 7.7 Rank-wise importance of symptom to fault of the drag system

Symptom	Fault weightage (rank)				
	BBF (rank)	DCP (rank)	SBP (rank)	ID (rank)	UF (rank)
CF	0.072 (5)	0.092 (6)	0.117 (6)	0.275 (1)	0.102 (6)
VF	0.073 (4)	0.224 (2)	0.097 (7)	0.026 (7)	0.033 (7)
VUS	0.565 (1)	0.114 (5)	0.158 (2)	0.072 (6)	0.153 (2)
LL	0.041 (7)	0.236 (1)	0.142 (3)	0.130 (4)	0.106 (5)
TF	0.101 (2)	0.060 (7)	0.134 (5)	0.199 (3)	0.121 (3)
Sp	0.094 (3)	0.114 (4)	0.138 (4)	0.240 (2)	0.110 (4)
US	0.054 (6)	0.159 (3)	0.214 (1)	0.058 (5)	0.375 (1)

In the relationship between symptom to fault analysis, the most observed symptom is the vibration and unwanted sound having weightage 56.5% (Table 7.7); therefore, it should be inspected first. Based on the observed symptom vibration and unwanted sound, the most influencing cause can be identified, and it is included in the preventive maintenance action plan (section 7.5.1). Similarly, when the fault 'insulation damage' is identified, it is observed that the most probable symptom is the current feedback having weightage 27.5%. Based on the symptom current feedback can be identified and defined the sequence of preventive maintenance action for cause (section 7.5.1). Similarly, when the fault drive control problem is identified, it is observed that the most probable symptom is lubrication level having weightage 23.6%. Based on the symptom lubrication level, the sequence of the preventive maintenance action plan for causes can be done. Similarly, when the fault drive control problem is identified, it is observed that the most likely symptom is the lubrication level having weightage 23.6%; therefore, if the symptom lubrication level is identified, the sequence of the proactive action plan can be made.

7.7 Summary

Two ANN models are used to make relationships between occurrence of cause, symptom, and fault of the drag system. The first ANN model develops the relationship between cause and symptom that can help to identify the sequence of responsible root causes to prevent the occurrence of fault when the symptom exceeds the threshold limit. The second ANN model develops the relationship between symptom to fault that can help to predict the possible occurrence of faults to avoid catastrophic failure of the dragline. The prediction accuracies of both the ANN models are very high, hence they provide reasonable confidence on prediction of faults and the subsequent development of maintenance policy. Therefore, the combination of both the ANN models can be used in real-time application to prepare the effective CBM policy to avoid the occurrence of fault as well as failure of the dragline.