Naïve Bayesian Classifier for On-line Remaining Useful Life Prediction of Degrading Bearings

Francesco Di Maio, Selina S. Y. Ng, Kwok-Leung Tsui, Enrico Zio

NAÏVE BAYSIAN CLASSIFIER FOR ON-LINE REMAINING USEFUL LIFE PREDICTION OF DEGRADING BEARINGS

Francesco Di Maio[1], Selina S. Y. Ng[2], Kwok Leung Tsui[2] and Enrico Zio[1,3]

[1] *Energy Department, Polytechnic of Milan*
Via Ponzio 34/3, 20133 Milano, Italy
f.dimaio@mail.polimi.it

[2] *Department of Manufacturing Engineering and Engineering Management (MEEM),*
City University of Hong Kong,
83 Tat Chee Avenue, Kowloon Tong, Hong Kong
sysng2@cityu.edu.hk kltsui@cityu.edu.hk

[3] *Ecole Centrale Paris and Supelec*
Grande Voie des Vignes, 92295 Chatenay-Malabry Cedex, France
enrico.zio@ecp.fr enrico.zio@supelec.fr enrico.zio@polimi.it

In this paper, the estimation of the Residual Useful Life (RUL) of degraded thrust ball bearings is made resorting to a data-driven stochastic approach that relies on an iterative Naïve Bayesian Classifier (NBC) for regression task. NBC is a simple stochastic classifier based on applying Bayes' theorem for posterior estimate updating. Indeed, the implemented iterative procedure allows for updating the RUL estimation based on new information collected by sensors located on the degrading bearing, and is suitable for an on-line monitoring of the component health status. The feasibility of the approach is shown with respect to real world vibration-based degradation data.

Key Words: Degradation, Prognostics, Condition Monitoring, Residual Useful Life, Naïve Bayesian Classifier, Thrust Ball Bearings.

1. Introduction

Bearings degradation and the prediction of their Residual Useful Life (RUL) are of great significance for many production and manufacturing industries. Indeed, according to [17], almost 40-50% of all rotating machinery failures are bearing related, thus greatly affecting productivity in those sectors in which rotating machinery plays a relevant role [10].

Maintenance of these components is currently performed under periodic schemes of inspections, which are not capable of giving due account to their actual health status. On the other hand, experience tells us that thrust bearing failures without warning can result in even catastrophic consequences [13, 20]. This has motivated a movement, in recent years, towards approaches for monitoring and predicting the trend of degradation of rotating machinery components within a philosophy of Condition-Based Maintenance (CBM) [22, 15, 16]. CBM allows prescheduled maintenance based on the information on the components RUL, reduces inventory costs of spare parts, reduces the risk of catastrophic failure and eliminates unexpected outage [24].

In practice, the estimation of the RUL may be difficult to obtain because the degradation state may not be directly observable and/or the vibration measurements may be affected by large noise and disturbances, as is the case of interest in our study.

From a methodological viewpoint, approaches for predicting rotating machinery failures can be grouped into two main categories [12]: traditional reliability approaches and condition-based approaches. Traditional reliability approaches are based on the distribution of event records of a population of identical units: their historical time-to-failure data are used to estimate the population characteristics, e.g., the mean time to failure. These methods are effective for units produced in high volumes, whereas they are not effective to characterize small numbers of component or even a single component currently running on a particular machine [12].

Condition-based approaches can be divided into two categories: physics-based models and data-based models [1]. Physics-based approaches attempt to set up comprehensive mathematical models to describe the physical phenomena underlying the components degradation process and failure modes, and to estimate their RUL. However, these models have been found inadequate in addressing the stochastic nature of defect-propagation; correspondingly, a number of stochastic models have been developed. For example, the Bayesian updating procedure for residual-life distributions estimation from component degradation signals proposed in [8] assumes a first model in which the degradation signal exhibits independent identically distributed (IID) errors for an exponential signal trajectory, and a second model in which the error fluctuations follow a Brownian Motion process (BM); Bayesian updating is used to improve the estimation of the stochastic parameters in the exponential model, thus eventually improving the estimate of the true signal trajectory. Both methods proposed in [8] are parametric and model-based.

Still, uncertainty due to assumptions and simplifications in the models may pose limitations to their applicability in practical industrial applications, where the failure-specific mechanistic knowledge is often hard to gather without interrupting operation [12].

In such cases, data-driven techniques may serve the purpose as they utilize monitored operational data related to system health and routinely collected from the machine, instead of building models based on comprehensive system physics and human expertise. They can be beneficial when understanding of first principles of system operation is not straightforward or when the system is so complex that developing an accurate model is prohibitively expensive.

In this work we resort to a non-parametric stochastic approach that relies on a Naïve Bayesian classifier (NBC) [5] to benefit both from the stochastic framework approach and the data-driven characteristic of the method; NBC consists in a simple stochastic classifier (based on applying Bayes' theorem) which iteratively allows for the on-line estimation of the RUL of a degrading thrust bearing by updating the estimates of the posterior, i.e., the RUL, based on the data collected from the sensors located on the monitored component which are compared to other data stored in a database and previously collected on degraded components of the same type. Despite its simplicity, the NBC learning scheme performs well on most classification tasks, and is often significantly more accurate than more sophisticated methods [7] also when applied for regression.

The paper contents are structured as follows. Section 2 presents the benchmark data of degrading thrust bearings taken from [8]. Section 3 contains the description of the approach at the basis of the RUL estimation, with an overview of the NBC framework. In Section 4, the results of the application of the approach here proposed are presented, and an evaluation of its performance is given. Section 5 shows a comparison of the results provided by the NBC-based approach and those given by the IID- and BM-based models [8]. Finally, some conclusions on the advantages and limitations of the approach here propounded are given in Section 6.

2. Case Study

The case study here illustrated has been analyzed previously with other model-based methods [8, 9, 10]. The database consists of degradation signals of 25 thrust ball bearings obtained from a run-to-failure experiment under accelerated testing conditions. The bearings are run under constant operating conditions and vibration frequency spectra are acquired every 2 minutes. Thus, the monitored signal consists of the average of the first seven harmonics of the amplitude of the bearing rotation frequency. Generally, the first part of the readings remains low, representing the normal operation phase where there are no cracks on the bearings. Then, after the failure onset, the vibration readings increase as defects develop and worsen, until the bearings is considered to be in a failure state and cannot be used, when a failure threshold d=0.03 V_{RMS} is reached [8], as shown in Figure 1. Among the 25 available degrading trajectories, 2 never reach the failure threshold, so that they are discarded. The performance of the approach will be compared with that obtained in [8], whose trajectories are shown in Figures 2-5, when a threshold equal to 2.5 mV_{RMS} is used to cut out the stable signal when the bearing is working in nominal conditions, i.e., in our analysis we will mainly concentrate on the second phase of the degradation.

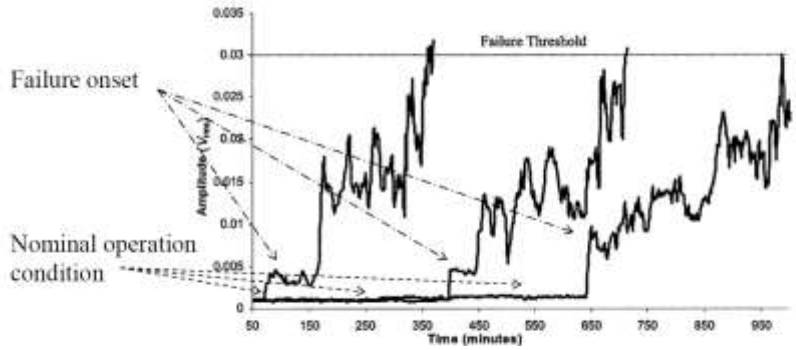

Figure 1. Typical behavior of degraded bearings measurements [8]

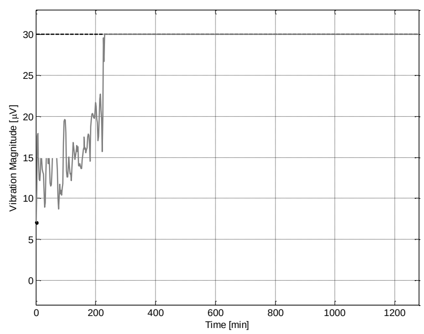

Figure 2. Degradation pattern 1

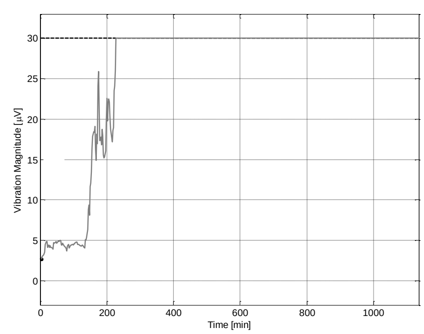

Figure 4. Degradation pattern 3

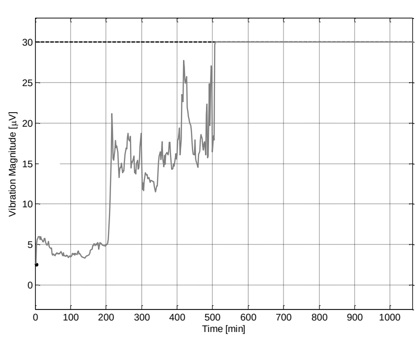

Figure 3. Degradation pattern 2

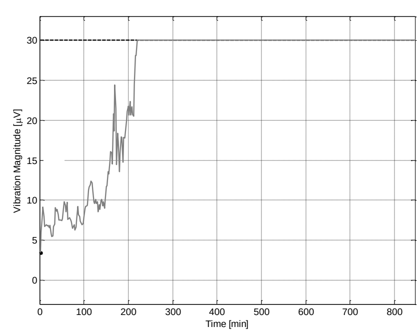

Figure 5. Degradation pattern 4

3. Methodology

Figure 6 shows a schematic sketch of the computational framework; for ease of illustration, a single signal $f(t)$ is considered (a generalization to the case of multidimensional trajectories is straightforward). It is continuously monitored throughout the time horizon of observation T, starting from (discrete) time $t=1$; at each inspection time T_j, $j=1,2,...,J$, its value is recorded and appended to the vector of the values collected at the time steps preceding T_j.

On the other hand, it is assumed that *N* trajectories of evolution in time of values of relevant signals (reference patterns) are available from measurements collected in dynamic degradation scenarios of the bearings under analysis. These trajectories last all the way to system failure, i.e., to the time when the signal reaches the threshold value beyond which the bearing loses its functionality. For reasons which will become clear in the following, the database constituting the reference pattern library is organized in a reference matrix $\overline{\overline{R}}_{[N\times(J+1)]}$, whose generic element $r(i,j)$ is the *j*-th segment of length $j+1$ of the values of the *i*-th reference pattern, $i=1,2,...,N$, $j=1,2,...,J$.

The idea underpinning the RUL estimation is to evaluate the similarity between the test trajectory pattern $f(t)$ of the developing degradation and the *N* reference trajectory patterns stored in the database through a NBC [7] and use the RULs of these latter to estimate, at each time T_j, the RUL of the former, accounting for how similar they are. In other words, the NBC calculates the probabilities of each feature given each class based on the set of reference trajectory patterns and uses these probabilities to calculate the probabilities of a test trajectory pattern belonging to each class given its feature values.

For the scope of this work, the implementation of the NBC went through three different steps: data pre-processing, training of the classifier, testing of the classifier. Data pre-processing entails the discretization of the target range, because it has been demonstrated that by so doing NBC for regression performs comparably to well known methods for time series prediction [14]. The pre-processing step, or data preparation, is a key step in the non-trivial Knowledge Discovery and Data Mining process upon which the success of the entire process depends [6, 4, 2, 3]. The second and the third steps are necessary within a supervised learning scheme, as it is the one proposed in this work.

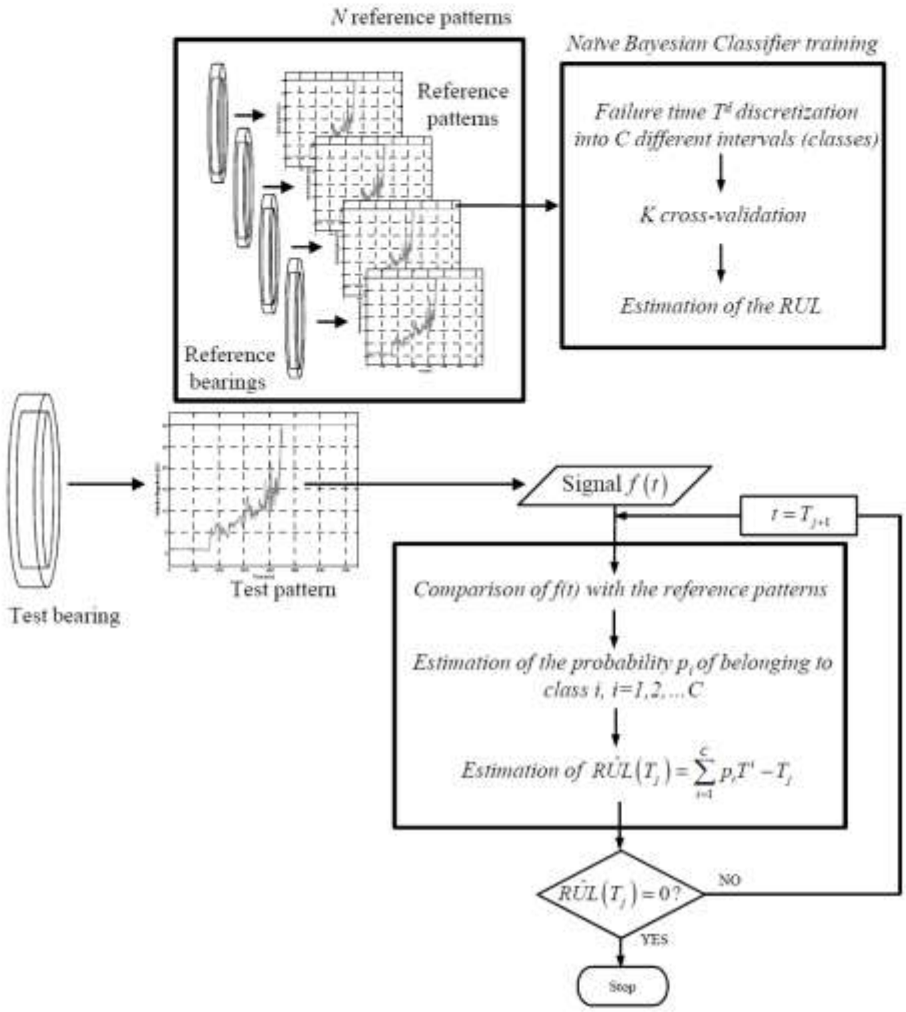

Figure 6. schematic sketch of the computational framework

a. Data pre-processing

When classifiers are used for regression, it is common to discretize the continuous target variable into a number of intervals [21, 23]. This decreases the number of classes to handle and increases class frequencies with the same number of training trajectories, so that to avoid over-fitting, improving the accuracy and the generalization capability of the learning scheme, and reducing at the same time the computational burden of the classification algorithms [23].

The time-point T^d at which each of the N=19 reference pattern hits the predetermined failure threshold d has been discretized by regression-by-discretization strategy [14]: the time-point T^d has been discretized into C equal width intervals of 40 min, each one with a reference mean life T^i, $i=1,2,\ldots,C,$ where $C <=N$, as the class label representative of the interval class label i, $i=1,2,\ldots C$ (Figure 7). To illustrate, discretization on this particular training and testing dataset (total 23 trajectories) reduces the number of class labels from 19 to 11.

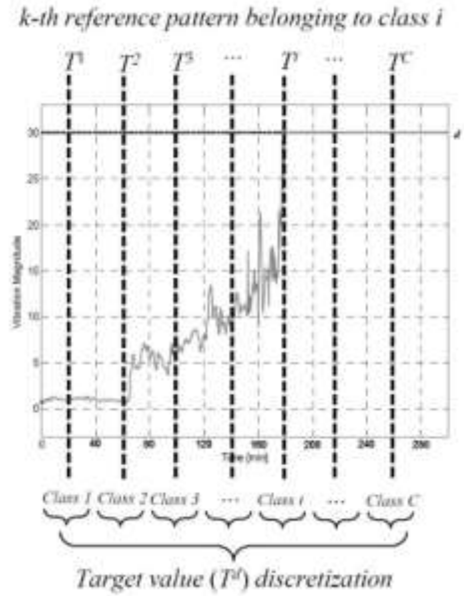

Figure 7. schematic sketch of the class discretization

b. Training of the NBC

A Bayesian classifier is a simple probabilistic classifier with strong (naïve) independence assumptions [5]. Although these assumptions make the model too simple to be realistic, Naïve Bayesian Classifiers generate much better than expected performances in some context [7, 18]. Moreover, NBC bear the advantage that they can be trained very efficiently in a supervised learning setting. By default, NBC handle nominal classes and assign the class with maximum probability to the instance.

Using Bayes' theorem, we can write the probability that a realization of feature variables $f_1, f_2, ..., f_J$, i.e. the vibration data at inspection times T_1, T_2,…, T_J, belong to class i as a conditional model:

$$p\left(i\middle| f_1,..., f_J\right) = \frac{p(i)\, p\left(f_1,..., f_J \middle| i\right)}{f\left(f_1,..., f_J\right)}$$

In practice we are only interested in the numerator of that fraction, since the denominator does not depend on i and the values of the features f_j are given, j=1,2,…,J , so that the denominator is effectively constant. The numerator is equivalent to the joint probability model which can be written as follows, using repeated applications of the definition of conditional probability:

$$p\left(C\middle| f_1,..., f_J\right) = p(i)\, p\left(f_1 \middle| i\right) p\left(f_2 \middle| i, f_1\right) ... p\left(f_J \middle| i, f_1, f_2, ..., f_{J-1}\right)$$

The "naïve" conditional independence assumes that each feature f_j is conditionally independent of every other feature f_t, $t \neq i$, so that:

$$p\left(f_j \middle| i, f_t\right) = p\left(f_j \middle| i\right)$$

Thus, the joint model can be expressed as:

$$p\left(i\middle| f_1,..., f_J\right) = p(i) \prod_{j=1}^{J} p\left(f_j \middle| i\right)$$

This means that under the above independence assumptions, the conditional distribution over the class variable i can be expressed like this:

$$p\left(i\middle| f_1,..., f_J\right) = \frac{1}{Z}\, p(i) \prod_{j=1}^{J} p\left(f_j \middle| i\right)$$

where Z (the evidence) is a scaling factor dependent only on $f_1, f_2, ..., f_J$, i.e., a constant if the values of the feature variables are known.

We resort to the supervised NBC classifier implemented in the Weka package [11] to perform the classification of the degrading trajectories of the thrust ball bearings presented in Section 2.

In this work, given the shortage of reference pattern, we proceed at training the NBC to provide $p\left(i\middle| f_1,..., f_J\right)$ with a 10-fold cross validation scheme [19] so that to improve the robustness of the predictions provided by the classifier. Reference patterns have been broken down into k=10 disjoint subsets of approximately equal size. Then, 10 experiments have been performed and for each experiment the k-th subset is removed: the system is trained on the remaining data. In other words, NBC is trained with 10-fold cross validation based on the N examples in $\overline{\overline{R}}_{[N,J+1]}$. Instead of testing on the held-out subset for each of the 10 folds, we evaluate the 10 NBC models on a separate set of m test trajectory patterns. Also, in contrast to the default NBC which chooses the nominal class with maximum probability, we utilize the probabilities of each of the possible class and the fact that the classes are continuous in our context. For each one of the 10 trained NBC, the estimated RUL is the weighted sum of all the class labels (time-to-failure) based on the probability of belonging to i-th class minus the time when the prediction is made, i.e., $R\hat{U}L_k\left(T_j\right)=\sum_{i=1}^{C} p_i T^i - T_j$. Finally, the estimated remaining useful life is taken equal to be the mean value of the k=10 RUL predictions provided by the 10 trained NBC, i.e., $R\hat{U}L\left(T_j\right)=\sum_{k=1}^{10} R\hat{U}L_k\left(T_j\right)/10$.

c. Testing of the NBC

For each m-th test trajectory, m=1,2,3,4, at each inspection time T_j, j=1,2,…,J , the vector of readings $f\left(t\right)$ is fed to the 10 trained NBC as input vector. Readings beyond the inspection time are set to be missing and are ignored by NBC. NBC estimates the probability of $f\left(t\right)$ belonging to each class i, i=1,2,…,C, the single RUL estimate $R\hat{U}L_k\left(T_j\right)$ is provided and the average RUL from the 10 predictions is obtained as $R\hat{U}L\left(T_j\right)$.

4. Results

Figures 8-11 are plots of the actual RUL (middle dotted line) and the predicted RUL (solid lines) for each test trajectory degrading patterns shown in Figures 2-5. Note that the RUL estimation is particularly stable and accurate throughout the life of degradation patterns 1, 3 and 4 (Figures 2, 4 and 5, respectively). In Figures 8-11, the ±10% error bounds are also give with respect to T^d , i.e., until T^d , the time at which the state hits the predetermined failure threshold d. Results indicate that the proposed method can give an high informative prediction at a sufficiently early stage of the component or structure life, so as to allow for preventive maintenance actions.

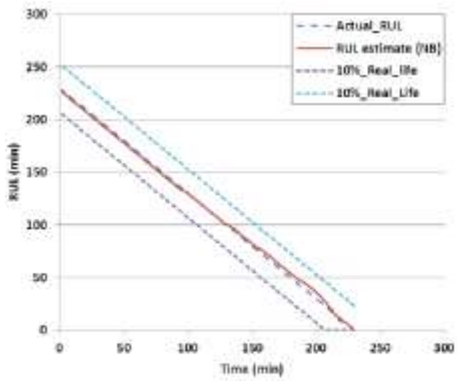

Figure 8. RUL estimation for the degrading pattern of Figure 2

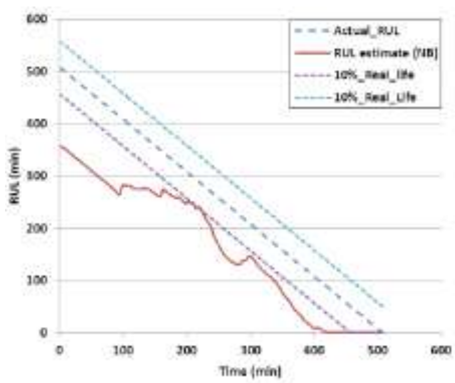

Figure 9. RUL estimation for the degrading pattern of Figure 3

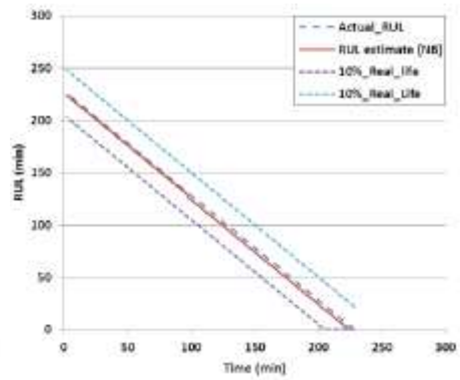

Figure 10. RUL estimation for the degrading pattern of Figure 4

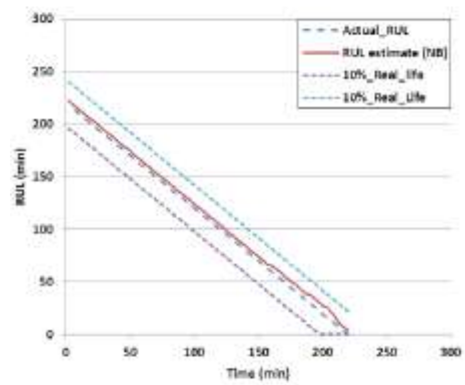

Figure 11. RUL estimation for the degrading pattern of Figure 5

The performance of the NBC classifier has been evaluated for each of the m=1,2,3,4 test pattern by the root mean square prediction error

$E_m = \sqrt{\sum_{j=1}^{J} \left(R\hat{U}L\left(T_j\right) - RUL\left(T_j\right) \right)^2 \Big/ T_m^d}$, where $RUL\left(T_j\right)$ is the actual RUL at inspection time T_j, $R\hat{U}L\left(T_j\right)$ is its estimate and T_m^d is the time at which the *m*-th test pattern hits the predetermined failure threshold *d*: the larger is the root mean square error, the smaller is the prediction accuracy. The root mean square error values are further scaled to the number of predictions for each test pattern for comparison with each other in Table 2. It is worth pointing out that the largest error occur in predicting the RUL of test pattern 2 (Figure 3): a further analysis has pointed out that this occurs because no reference patterns are stored in the database with a life equal to T_2^d , whereas NBC performs much better with degradation patterns whose behavior is well covered by the reference patterns stored in the training dataset.

Deagrading pattern (testing trajectory)					
1	**2**	**3**	**4**	**Average**	**Variance**
0.024	0.378	0.029	0.044	0.119	0.030

Table 2. Evaluation results – root mean square prediction error

5. Comparison with other methods

The results published for exponential Brownian Motion error exponential process (BM) and exponential stochastic process with Independent Identically Distributed error (IID) model in [8] are compared with that of the NBC at times equal to 0,5, 0.75 and 0.9 of the bearing failure time T^d. The best prediction among BM, IID and NBC for each of the four test bearings is highlighted in bold in Table 3. Note that these do not serve as direct comparisons because the data for training the NBC in this work is a subset, rather than an identical set, of that used for estimating the prior distributions in [8].

It is important to note that NBC always generates much more useful predictions at $0.5T^d$ for all test bearings rather than the other two methods. Generally, NBC performs better at early stage of the failure onset, i.e., RUL estimates are much more accurate than those provided by BM and IID models, and when training patterns with similar life-span exists, e.g., bearings 1, 3 and 4. When data is not available for training, either IID or BM models perform much better at earlier and later stage than NBC, as it occurs for bearing 2. This means that if training data is not available and good RUL predictions are only required at later stages e.g., the last hour before reaching the failure threshold, BM and IID have to be preferred. The use of both methods in parallel (hybridization of model-based and data-based approaches, i.e., BM and IID, and NBC, respectively) could provide even better estimation of RUL of thrust ball bearings. More work can be done in this path, and some efforts can be devoted to the reduction of the dimension of data set needed for training NBC while maintaining the capability of RUL prediction.

Bearing	**Actual life**	**Prediction time**	**IID**			**BM**			**NBC**		
1	**230**	$\boldsymbol{T^d}$	**5%**	**50%**	**95%**	**5%**	**50%**	**95%**	**5%**	**Avg.**	**95%**
	116	**0.5** $\boldsymbol{T^d}$	360	732	>10000	209	411	1797	206	**230**	288
	174	**0.75** $\boldsymbol{T^d}$	226	356	774	187	251	811	206	**234**	288
	208	**0.9** $\boldsymbol{T^d}$	290	290	514	213	247	731	206	**234**	288
2	**508**	$\boldsymbol{T^d}$	**5%**	**50%**	**95%**	**5%**	**50%**	**95%**	**5%**	**Avg.**	**95%**
	254	**0.5** $\boldsymbol{T^d}$	284	**562**	>10000	270	346	1062	326	414	492
	382	**0.75** $\boldsymbol{T^d}$	402	774	>10000	399	485	1461	406	**408**	408
	458	**0.9** $\boldsymbol{T^d}$	486	1166	>10000	479	**589**	1861	406	418	448
3	**228**	$\boldsymbol{T^d}$	**5%**	**50%**	**95%**	**5%**	**50%**	**95%**	**5%**	**Avg.**	**95%**
	114	**0.5** $\boldsymbol{T^d}$	168	502	>10000	142	254	1656	206	**224**	288
	172	**0.75** $\boldsymbol{T^d}$	202	688	>10000	185	263	1475	206	**224**	288
	206	**0.9** $\boldsymbol{T^d}$	226	630	>10000	221	305	1703	206	**224**	288
4	**220**	$\boldsymbol{T^d}$	**5%**	**50%**	**95%**	**5%**	**50%**	**95%**	**5%**	**Avg.**	**95%**
	110	**0.5** $\boldsymbol{T^d}$	238	524	>10000	144	248	1034	206	**224**	288
	166	**0.75** $\boldsymbol{T^d}$	378	378	8166	173	215	831	206	**226**	288
	198	**0.9** $\boldsymbol{T^d}$	338	338	2198	204	248	830	206	**228**	288

Table 3. Failure times evaluated at three prediction intervals using BM, iid, RVM and NB for four degrading patterns

6. Conclusions

In this paper, a data-driven stochastic approach that relies on an iterative Naïve Bayesian Classifier (NBC) is presented as a feasible tool for estimating the Residual Useful Life (RUL) of degraded thrust ball bearings. NBC is fed with the vibration signal collected on the component and provide the RUL of the most similar degradation trajectory among a set of reference patterns that are stored in a database. Indeed, the simple Bayesian framework upon which NBCs are founded allows for an iterative updating procedure for updating the RUL estimation based on new information collected by sensors located on the degrading bearing. The results show that the approach is suitable for an on-line monitoring of component health status in a real industrial application.

References

[1] Chiang, L. H., Russel, E., Braatz, R., "Fault detection and diagnosis in industrial systems", Springer-Verlag, London, 2001.

[2] Cios, K. J. and Kurgan, L. A.. "Trends in data mining and knowledge discovery". In In: Pal N.R., Jain, L.C. and Teoderesku, N. (Eds.), Knowledge Discovery in Advanced Information Systems, pages 200–2. Springer, 2002

[3] CRISP-DM. Cross industry standard process for data mining. http://www.crisp-dm.org/Process/index.htm, accessed on 01 Apr 2011.

[4] Debuse, J., de la Iglesia, B., Howard, C., and Rayward-Smith, V. "A Methodology for Knowledge Discovery: a KDD Roadmap". Sys-c99-01, School of Computing Sciences, University of East Anglia, 1999.

[5] Duda, R.O., and Hart, P.E. "Bayes Decision Theory". Chapter 2 in Pattern Classification and SceneAnalysis, pp. 10–43. John Wiley, 1973.

[6] Fayyad, U.M., Piatetsky-Shapiro, G., and Smyth, P., "Knowledge Discovery and Data Mining: Towards a Unifying Framework", Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD96), Portland, OR. AAAI Press, 1996

[7] Frank, E. and Trigg, L. and Holmes, G. and Witten, I.H., "Technical note: Naive Bayes for regression", Machine Learning, Vol. 41, No. 1, 2000.

[8] Gebraeel, N. G., Lawley, M.A., Li. R. , Ryan, J.K. "Residual.-life distributions from component degradation signals: a Bayesian approach", IIE Transactions, 37, pp. 543-557, 2005.

[9] Gebraeel, N.,G., "Sensory- Updated Residual Life Distributions for Components with Exponential Degradation Patterns, IEEE Transanctions on automatic science and engineering, Vol. 3, No. 4, 2006.

[10] Gebraeel, N. G. , Elwany, A. , Pan, J. , "Residual Life Predictions in the absence of prior degradation Knowledge", IEEE Transaction on Reliability, Vol. 58, No. 1, 2009.

[11] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I. H., "The WEKA Data Mining Software: An Update"; SIGKDD Explorations, Volume 11, Issue 1, 2009.

[12] Heng, A., Zhang, S., Tan, A.C.C., Mathew, J., "Rotating machinery prognostics: State of the art, challenges and opportunities", Mechanical Systems and Signal Processing 23 (2009) 724–739.

[13] Kelly, A., "Maintenance and its management", Proceedings of the Communication Conference, Surrey, England, 1989.

[14] Kononenko, I., “Semi-naive Bayesian classifier”. In Proceedings of the 6th European Working Session on Learning, Porto, Portugal (pp. 206–219). Berlin: Springer-Verlag.

[15] Mann, L. Saxena, A., Knapp G.M., “Statistical-based or condition-based preventive maintenance”, Journal of Quality in Maintenance Engineering, Vol. 1 No. 1, 45-59, 1995.

[16] Marseguerra, M., Zio, E., Podofillini, L., “Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation”, Reliab. Eng. System Safety, 77:151-65, 2002.

[17] Nandi, S., Toliyat, H.A., Li, X., “Condition Monitoring and Fault Diagnosis of Electrical Motors – A Review”, IEEE Transaction on Energy Conversion, Vol. 20, No. 4, 2005.

[18] Sahami, M., Dumais, S., Heckerman, D., and Horvitz, E. (1998), “A Bayesian approach to filtering junk e-mail”, In AAAI-98 Workshop on Learning for Text Categorization, volume 460.

[19] Salzberg, S., “On comparing classifiers: Pitfalls to avoid and a recommended approach“ Data Mining and Knowledge Discovery, 1(3):317-328. 1997.

[20] Singh, G.K., Saleh Al Kazzazb, S.A., “Induction machine drive condition monitoring and diagnostic research—a survey”, Electric Power Systems Research Volume 64, Issue 2, February 2003, Pages 145-158.

[21] Torgo, L., Gama, J., “Regression using classification algorithms”, Intelligent Data Analysis, Vol. 1, No. 1-4, pp. 275-292, 1997.

[22] Williams, J. H., Davies, A., Drake, I. R., “Condition-based maintenance and machine diagnostics”, Chapman and Hall, ISBN 0 4 I2465 00 0, I87pp, 1994.

[23] Yang, Y., Webb, G.I., “Discretization for naive-Bayes learning: Managing discretization bias and variance”, Machine learning, Vol. 74, No. 1, 2009.

[24] Yam, RCM and Tse, PW and Li, L. and Tu, P., “Intelligent predictive decision support system for condition-based maintenance”, The International Journal of Advanced Manufacturing Technology, Vol. 17, No. 5, 2001.