Anomaly Detection for the Prediction of Ultimate Tensile Strength in Iron Casting Production

Igor Santos, Javier Nieves, Xabier Ugarte-Pedrero, and Pablo G. Bringas

S[3]Lab, DeustoTech - Computing, Deusto Institute of Technology
University of Deusto,
Avenida de las Universidades 24, 48007
Bilbao, Spain
{isantos,jnieves,xabier.ugarte,pablo.garcia.bringas}@deusto.es

Abstract. Mechanical properties are the attributes that measure the faculty of a metal to withstand several loads and tensions. In particular, ultimate tensile strength is the force a material can resist until it breaks. This property is one of the variables to control in the foundry process. The only way to examine this feature is to apply destructive inspections that make the casting invalid with the subsequent cost increment. Modelling the foundry process as an expert knowledge cloud allows machine-learning algorithms to forecast the value of a certain variable; in this case, the probability of a certain value of ultimate tensile strength for a foundry casting. Nevertheless, this approach needs to label every instance in the training dataset for generating the model that can foresee the value of ultimate tensile strength. In this paper, we present a new approach for detecting castings with an invalid ultimate tensile strength value based on anomaly detection methods. This approach represents correct castings as feature vectors of information extracted from the foundry process. A casting is then classified as correct or not correct by measuring its deviation to the representation of normality (correct castings). We show that this method is able to reduce the cost and the time of the tests currently used in foundries.

Keywords: Fault prediction, anomaly detection, ultimate tensile strength, industrial processes optimisation.

1 Introduction

Foundry can be considered as one of the axis of current economy because a huge number of castings are manufactured to be part of more complex systems e.g., the brake component of a car, the propeller of a boat, components of the wings of an aircraft and the trigger in a weapon. Therefore, if one of the pieces is faulty, it can be detrimental to both individuals and businesses activities.

Unfortunately, although there are many standards and methods to check the quality of the produced castings, these are performed once the manufacturing of

A. Hameurlain et al. (Eds.): DEXA 2011, Part II, LNCS 6861, pp. 519–526, 2011.

the casting has been completed. Most of the used techniques for the assurance of failure-free foundry processes are exhaustive production control and diverse simulation techniques [1] but they are extremely expensive. In this paper, we focus on the so-called Ultimate Tensile Strength (UTS). This mechanical property is defined as the force a casting can resist until it breaks; i.e., the maximum stress any material can withstand when subjected to tension. Manufactured iron castings must assure certain value (or threshold) of UTS to pass the strict quality tests. Unfortunately, the only available approach to examine the UTS breaks the piece, incurring in a cost increment within the process.

In our previous work [2–4], we have proven the ability of several machine-learning classifiers for the prediction of mechanical properties. We used Bayesian networks, support vector machines, decision trees, artificial neural networks and support vector machines, to identify the best overall machine-learning classifier capable of predicting the value of UTS and to reduce the noise in the manual data-gathering process [5]. However, machine-learning classifiers (or supervised learning methods) require a high number of labelled castings for each of the classes (i.e., faulty and not-faulty castings) to train the different models. However, it is quite difficult to acquire this amount of labelled data for a real-world problem such as production control. To generate these data, a time-consuming process of analysis is mandatory that renders in a cost increment during the process.

Given this background, we present here a method to classify castings to foresee the value of UTS that it is based on anomaly detection methods. This approach is able to determine whether a casting contains a valid UTS value or not by comparing several features extracted from the production of the castings with a dataset composed only of the features of valid castings. Therefore, if the casting under prediction presents a considerable deviation to what is considered as normal (the previously stored correct castings), the casting is considered to be faulty and, thus, there may be a high probability that the casting has an invalid value of UTS. This method deals with the aforementioned problem, achieving a reduction in the required number of castings to be labelled.

Summarising, our main contributions are: (i) we select a set of variables extracted from the foundry process to determine whether a casting has a valid UTS value or not and provide a relevance measure for each variable based on information gain, (ii) we propose an anomaly-detection-based architecture for UTS prediction, by means of weighted comparison against a dataset composed of only correct castings and (iii) we evaluate the method using three different deviation measures and show that this method can reduce the need of labelling castings.

2 Foundry Processes and Mechanical Properties

Several factors, for instance the extreme conditions in which it is performed, make the foundry process very complex. Starting from the raw material to the manufactured item, this procedure involves numerous stages, several of which may be performed in parallel. When it comes to iron ductile castings, this process presents the following phases:

- **Pattern making:** In this phase, the moulds (exteriors) and the kernels (interiors) are produced in wood, metal or resin to be used in the generation of the moulds where the final casting will be built.
- **Mould and kernel generation:** Although other methods exist, the sand mould is most widespread method for the manufacturing of foundry castings. The sand is mixed with clay, water or other chemical binders. Next, specialised machines generate the two halves of the mould and they unite them in order to create the container where the melt metal will be introduced.
- **Melting and pouring:** The raw metals are melt, mixed and poured onto the sand moulds.
- **Cooling:** The solidification of the castings is controlled in the cooling lines until this process is finished.
- **Finishing:** In order to finish the process, once the casting is cleaned, some actions are usually performed like thermal treatment, ratification of defects in welds and so on.

Once these phases finish, foundry materials are subject to forces (loads). Engineers calculate these forces and how the material deforms or breaks as a function of applied load, time or other conditions. It is important to know how mechanical properties affect to iron castings [6], because they directly affect the final quality of the manufactured casting. The most important mechanical properties of foundry materials are the following ones [7]: strength, hardness, resilience, elasticity, plasticity, brittleness, ductility and malleability. In this work, we focus on Ultimate Tensile Strength (UTS) that is a type of strength, which is the property that enables a metal to resist deformation under load. The testing method of UTS is conducted as follows. First, a scientist prepares a testing specimen from the original casting. Second, the specimen is placed on the tensile testing machine. Finally, this machine pulls the sample from both ends and measures the force required to break the specimen apart and how much the sample stretches before breaking.

The complexity of UTS prediction of the resulting castings arises mainly from the large number of variables involved in the production process and, therefore, this variables influence the final design of castings. The total number of variables we focus on has been reduced to 24, and more specifically, the control variables can be divided into metal-related variables and variables related to the mould.

- **Metal-related**
  - *Composition:* type of treatment, inoculation and quantities.
  - *Thermal:* Nucleation potential and quality of the mixture, obtained by thermal analysis [8].
  - *Pouring:* Pouring duration and temperature.

– **Mould-related**
  - Sand: types of additives used for sand, the specific characteristics of the sand.
  - Mould: mould and machine parameters used.

Generally, the size and geometry of the casting play a very important and, therefore, we also included several variables to monitor these features. Similarly, the system takes into account the parameters related to the configuration of each machine working in the manufacturing process. Also, we added other variables such as cooling rate and heat treatment applied to the piece.

Although we have already obtained overall good results using a machine-learning-based approach for predicting imperfections and mechanical properties [2, 4, 5, 9–13], these approaches require a manual labour to label every instance of the training dataset. This process can be specially time-consuming and, also, means a cost increment.

We present here an anomaly-based approach that only requires labelling the correct castings and that measures the deviations of the inspected pieces with these previous stored castings. Such an approach will reduce the efforts of labelling castings, working with less information available. To this end, as we mentioned before, we manage 24 variables extracted from the foundry process. To provide a more accurate deviation measure, we apply relevance weights to each feature based on Information Gain (IG) [14]. This is done because IG provides a ratio for each characteristic that measures its importance to consider if a casting is valid or not. These weights were calculated from a real dataset acquired from a foundry specialised in safety and precision components for the automotive industry. The dataset is composed of 645 correct castings and 244 faulty castings.

3 Anomaly Detection

Through the features described in the previous section, our method represents valid castings as points in the feature space. When a casting is being inspected our method starts by computing the values of the point in the feature space. This point is then compared with the previously calculated points of the valid foundry castings.

To this end, distance measures are required. In this study, we have used the following distance measures:

– **Manhattan Distance:** This distance between two points v and u is the sum of the lengths of the projections of the line segment between the points onto the coordinate axes: $d(x, i) = \sum_{i=0}^{n} |x_i - y_i|$ where x is the first point; y is the second point; and x_i and y_i are the ith component of first and second point, respectively.
– **Euclidean Distance:** This distance is the length of the line segment connecting two points. It is calculated as: $d(x, y) = \sum_{i=0}^{n} \sqrt{v_i^2 - u_i^2}$ where x is the first point; y is the second point; and x_i and y_i are the ith component of first and second point, respectively.

- **Cosine Similarity:** It is a measure of similarity between two vectors by finding the cosine of the angle between them [15]. Since we are measuring distance and not similarity we have used $1-\textit{Cosine Similarity}$ as a distance measure: $d(x,y) = 1 - \cos(\theta) = 1 - (\boldsymbol{v} \cdot \boldsymbol{u})/(||\boldsymbol{v}|| \cdot ||\boldsymbol{u}||)$where $\boldsymbol{v}$ is the vector from the origin of the feature space to the first point x, $\boldsymbol{u}$ is the vector from the origin of the feature space to the second point y, $\boldsymbol{v} \cdot \boldsymbol{u}$ is the inner product of $\boldsymbol{v}$ and $\boldsymbol{u}$. $||\boldsymbol{v}|| \cdot ||\boldsymbol{u}||$ is the cross product of $\boldsymbol{v}$ and $\boldsymbol{u}$. This distance ranges from 0 to 1, where 1 means that the two evidences are completely different and 0 means that the evidences are the same (i.e., the vectors are orthogonal between them).

Using these measures, we are capable of computing the deviation of a casting respect to a set of not faulty castings. Since we have to compute this measure with all the points representing valid castings, a combination metric is required in order to obtain a final value of distance which considers every measure performed. To this end, our system employs 3 very simplistic rules: (i) select the mean value, (ii) select the lowest distance value and (iii) select the highest value of the computed distances. In this way, when our method inspects a casting a final distance value is acquired, which will depend on both the distance measure and the combination rule.

4 Empirical Validation

In order to evaluate our anomaly-based faulty casting detector, we collected a dataset from a foundry, which is specialised in safety and precisions components for the automotive industry, principally in disk-brake support with a production over 45,000 tons a year.

The acceptance/rejection criterion of the studied models resembles the one applied by the final requirements of the customer. Pieces flawed with an invalid UTS must be rejected due to the very restrictive quality standards (which is an imposed practice by the automotive industry). To this extent, we have defined two risk levels: *Valid* (more than 370 MPa) and *Invalid* (less than 370 MPa).

We worked with two different references, in other words, type of pieces and, in order to test the proposed method, with the results of the destructive inspections of the 889 production stocks performed in beforehand. More accurately, the dataset comprises 645 correct castings and 244 faulty castings.

Specifically, we conducted the next configuration for the empirical validation:

1. **Cross validation.** Despite the small dataset, we have to use as much of the available information in order to obtain a proper representation of the data. To this extent, we performed a 5-fold cross-validation [16] over the correct castings to divide it into 5 different divisions of 552 castings for representing normality and 138 for testing. In this way, each fold is composed of 516 not faulty castings that will be used as representation of normality and 373 testing castings, from which 129 are valid castings and 244 are faulty castings.

2. **Calculating distances and combination rules.** We extracted the aforementioned characteristics and employed the 3 different measures and the 3 different combination rules described in Section 3 to obtain a final measure of deviation for each testing evidence. More accurately, we applied the following distances: (i) Manhattan Distance, (ii) Euclidean Distance and (iii) Cosine Similarity. For the combination rules we have tested the followings: (i) mean value, (ii) lowest distance and (iii) highest value.
3. **Defining thresholds.** For each measure and combination rule, we established 10 different thresholds to determine whether a casting is valid or not.
4. **Testing the method.** We evaluated the accuracy of the proposed model by measuring False Negative Rate (FNR) and False Positive Rate (FPR). In particular, FNR is defined as: $FNR(\beta) = FN/(FN + TP)$ where TP is the number of faulty castings correctly classified (true positives) and FN is the number of faulty castings misclassified as valid castings (false negatives). On the other hand, FPR is defined as: $FPR(\alpha) = FP/(FP + TN)$ where FP is the number of valid castings incorrectly detected as faulty castings while TN is the number of valid castings correctly classified.

Table 1. Results for different combination rules and distance measures. The results in bold are the best for each combination rule and distance measure. Our method is able to detect more than 78 % of the faulty castings although with a FPR higher than 40 %.

	$1 - Cosine\ Similarity$			$EuclideanDistance$			$ManhattanDistance$		
Combination	**Threshold**	**FNR**	**FPR**	**Threshold**	**FNR**	**FPR**	**Threshold**	**FNR**	**FPR**
Mean	0.2038	0.000	0.987	0.110	0.000	1.000	0.225	0.000	0.994
	0.2413	0.119	0.688	0.1184	0.087	0.825	0.2513	0.093	0.831
	0.2789	0.227	0.530	0.1265	0.129	0.664	0.2778	0.252	0.459
	0.3164	**0.566**	**0.189**	**0.1347**	**0.215**	**0.426**	**0.3043**	**0.520**	**0.190**
	0.3539	0.757	0.128	0.1428	0.551	0.192	0.3307	0.741	0.116
	0.3914	0.846	0.079	0.1510	0.724	0.114	0.3572	0.846	0.069
	0.4290	0.917	0.051	0.1591	0.831	0.062	0.3837	0.906	0.049
	0.4665	0.951	0.029	0.1673	0.931	0.029	0.4102	0.950	0.027
	0.5040	0.963	0.021	0.1754	0.972	0.006	0.4367	0.972	0.006
	0.5415	0.984	0.000	0.1836	0.999	0.000	0.4632	1.000	0.000
Maximum	0.5220	0.000	0.986	0.1810	0.000	0.978	0.4734	0.000	0.964
	0.5716	0.027	0.888	0.1887	0.081	0.899	**0.4984**	**0.041**	**0.891**
	0.6212	**0.063**	**0.818**	0.1964	0.170	0.810	0.5234	0.267	0.782
	0.6708	0.167	0.765	0.2042	0.238	0.686	0.5484	0.530	0.649
	0.7204	0.287	0.672	**0.2119**	**0.299**	**0.613**	0.5734	0.722	0.488
	0.7700	0.429	0.527	0.2196	0.521	0.471	0.5984	0.845	0.333
	0.8197	0.677	0.424	0.2274	0.733	0.355	0.6233	0.897	0.193
	0.8693	0.865	0.283	0.2351	0.841	0.201	0.6483	0.936	0.094
	0.9189	0.957	0.075	0.2428	0.964	0.105	0.6733	0.973	0.023
	0.9685	0.993	0.000	0.2505	0.989	0.000	0.6983	0.992	0.000
Minimum	0.0000	0.00000	0.94884	0.0000	0.000	0.978	0.0000	0.000	0.978
	0.0212	**0.804**	**0.060**	**0.0093**	**0.326**	**0.418**	**0.0169**	**0.320**	**0.415**
	0.0424	0.931	0.021	0.0186	0.559	0.248	0.0338	0.599	0.223
	0.0636	0.977	0.012	0.0279	0.727	0.145	0.0507	0.755	0.116
	0.0849	0.997	0.004	0.0371	0.809	0.046	0.0675	0.843	0.037
	0.1061	1.000	0.001	0.0464	0.879	0.021	0.0844	0.909	0.013
	0.1273	1.000	0.001	0.0557	0.985	0.004	0.1013	0.965	0.006
	0.1485	1.000	0.001	0.0650	0.999	0.004	0.1182	0.998	0.001
	0.1697	1.000	0.001	0.0743	0.999	0.003	0.99918	0.001	0.001
	0.1910	1.000	0.000	0.0836	1.000	0.000	0.1520	1.000	0.000

Table 1 shows the obtained results. Euclidean and Manhattan distances, despite of consuming less processing time, have achieved better results than cosine-similarity based distance for the tested threshold. Our anomaly-based faulty casting detector, for each distance measure, accomplished its best results selecting the mean value for computing the deviation of a casting respect to the not faulty castings. In particular, our detector is able to detect more than 78 % of faulty castings (using Euclidean distance), maintaining the rate of misclassified correct castings in 42.6%. Nevertheless, all distances obtain similar results.

5 Conclusions

Foreseeing the value of UTS in ductile iron castings is one of the most hard challenges in foundry-related research. Our work in [2, 4] pioneered the application of artificial intelligence methods to the prediction of the value of UTS.

This time, our main contribution is the anomaly-detection-based approach employed for UTS prediction. In contrast to our previous approaches, this method only need previously labelled the correct castings and it measures the deviation of castings respect to normality (castings with a valid value of UTS). Although anomaly detection systems tend to produce high error rates, in our case, the criteria establishes that a high false positive rate is tolerable whereas a high false negative rate is not. Therefore, our method is suitable for its direct application within real foundries.

Anyway, it presents some limitations that should be studied in further work. Firstly, we cannot identify different levels of warnings as we did in our previous works. In this case, we only can classify the castings as correct or faulty. Nevertheless, we could compute it using another anomaly detection techniques such as clustering based or nearest neighbour based anomaly detection.

Secondly, this kind of techniques based on the measurement of distances cannot achieve good results if the training data is disperse. In other words, if the normality cannot be represented as a compact group of instances, the threshold that allows to split the evidences between correct and faulty does not adjust to have its best behaviour. Nevertheless, this fact is solved due to the nature of the productions process, since all the castings are always produced in similar way. Hence, generated vectors of castings are close between each other, representing the normality in a good way in order to measure the distances between correct and faulty castings.

Finally, it is important to consider efficiency and processing time. Our system compares each casting against a relative big dataset (244 vectors for each fold). Despite Euclidean and Manhattan distances are easy to compute, cosine distance and more complex distance measures such as Mahalanobis distance may take too much time to process every casting under analysis. For this reason, in further work we will emphasise on improving the system efficiency by reducing the whole dataset to a limited amount of samples which is sufficiently representative.

References

1. Sertucha, J., Loizaga, A., Suárez, R.: Improvement opportunities for simulation tools. In: Proceedings of the 16th European Conference and Exhibition on Digital Simulation for Virtual Engineering (2006) (invited talk)
2. Nieves, J., Santos, I., Penya, Y.K., Rojas, S., Salazar, M., Bringas, P.G.: Mechanical properties prediction in high-precision foundry production. In: Proceedings of the 7^{th} IEEE International Conference on Industrial Informatics (INDIN 2009), pp. 31–36 (2009)
3. Nieves, J., Santos, I., Penya, Y.K., Brezo, F., Bringas, P.G.: Enhanced foundry production control. In: Bringas, P.G., Hameurlain, A., Quirchmayr, G. (eds.) DEXA 2010. LNCS, vol. 6261, pp. 213–220. Springer, Heidelberg (2010)
4. Santos, I., Nieves, J., Penya, Y.K., Bringas, P.G.: Machine-learning-based mechanical properties prediction in foundry production. In: Proceedings of ICROS-SICE International Joint Conference (ICCAS-SICE), pp. 4536–4541 (2009)
5. Nieves, J., Santos, I., Bringas, P.: Overcoming data gathering errors for the prediction of mechanical properties on high precision foundries. In: World Automation Congress (WAC), pp. 1–6. IEEE, Los Alamitos (2010)
6. Gonzaga-Cinco, R., Fernández-Carrasquilla, J.: Mecanical properties dependency on chemical composition of spheroidal graphite cast iron. Revista de Metalurgia 42, 91–102 (2006)
7. Lung, C.W., March, N.H.: Mechanical Properties of Metals: Atomistic and Fractal Continuum Approaches. World Scientific Pub. Co. Inc., Singapore (1992)
8. Larrañaga, P., Sertucha, J., Suárez, R.: Análisis del proceso de solidificación en fundiciones grafíticas esferoidales. Revista de Metalurgia 42(4), 244–255 (2006)
9. Santos, I., Nieves, J., Penya, Y.K., Bringas, P.G.: Optimising machine-learning-based fault prediction in foundry production. In: Omatu, S., Rocha, M.P., Bravo, J., Fernández, F., Corchado, E., Bustillo, A., Corchado, J.M. (eds.) IWANN 2009. LNCS, vol. 5518, pp. 554–561. Springer, Heidelberg (2009)
10. Santos, I., Nieves, J., Bringas, P., Penya, Y.: Machine-learning-based defect prediction in highprecision foundry production. In: Becker, L.M. (ed.) Structural Steel and Castings: Shapes and Standards, Properties and Applications, pp. 259–276. Nova Publishers (2010)
11. Nieves, J., Santos, I., Penya, Y.K., Brezo, F., Bringas, P.G.: Enhanced foundry production control. In: Bringas, P.G., Hameurlain, A., Quirchmayr, G. (eds.) DEXA 2010. LNCS, vol. 6261, pp. 213–220. Springer, Heidelberg (2010)
12. Santos, I., Nieves, J., Bringas, P.: Enhancing fault prediction on automatic foundry processes. In: World Automation Congress (WAC), pp. 1–6. IEEE, Los Alamitos (2010)
13. Santos, I., Nieves, J., Penya, Y.K., Bringas, P.G.: Towards noise and error reduction on foundry data gathering processes. In: Proceedings of the International Symposium on Industrial Electronics (ISIE), pp. 1765–1770 (2010)
14. Kent, J.: Information gain and a general measure of correlation. Biometrika 70(1), 163–173 (1983)
15. Tata, S., Patel, J.: Estimating the Selectivity of tf-idf based Cosine Similarity Predicates. SIGMOD Record 36(2), 75–80 (2007)
16. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the International Joint Conference on Artificial Intelligence, vol. 14, pp. 1137–1145 (1995)