Journal of Materials Science manuscript No.
(will be inserted by the editor)

Supporting Information: Development of Vickers hardness prediction models via microstructural analysis and machine learning

Sucheta Swetlana · **Nikhil Khatavkar** · **Abhishek Kumar Singh**

Received: date / Accepted: date

Data Collection

Table S1 includes the references of the experimental reports used as input data for development of ML model.

Table S1: References for the microstructures and experimental data used in this work.

Sl.No.	System	Reference
1	NiCrAlNb	[1]
2	CoNiAlTi	[2]
3	CoNiAlWTi	[2]
4	CoAlWTa	[3]
5	CoAlWMo	[3]
6	CoAlWCrNi	[4]
7	CoAlWCr	[4]
8	CoAlW	[4]
9	CoAlWMo	[4]
10	CoAlWV	[4]
11	CoAlWTi	[4]
12	CoAlWTa	[4]
13	CoAlWNi	[4]
14	CoAlWNiV	[4]
15	CoAlWSi	[4]
16	CoAlWFe	[4]
17	NiAlTiV	[5]
18	NiAlNbV	[6]
19	NiCrAlCoWTaTiMoC	[7]
20	NiCrCoMoWTiAlCBZr	[8]
21	NiCrCoMoAlTiTaZrCB	[9]
22	NiCrCoTiAlMoHfW	[10]
23	NiFeCrMoTiNbAlCB	[11]
24	NiCrCoMoWTiAlCBZr	[12]
25	NiTiAlNbCrCoFeMoWC	[13]
26	CoNiTiCrMoAlW	[14]
27	NiCrMoFeWMnSiCB	[15]
28	NiCrMoTiAlNbCo	[16]
29	CoAlWTaTiCrNiHfC	[17]

Sucheta Swetlana
Materials Research Centre, Indian Institute of Science, Bangalore, India, 560012 E-mail: suchetas@iisc.ac.in

Nikhil Khatavkar
Materials Research Centre, Indian Institute of Science, Bangalore, India, 560012 E-mail: nikhilk@iisc.ac.in

Abhishek Kumar Singh
Materials Research Centre, Indian Institute of Science, Bangalore, India, 560012 E-mail: abhishek@iisc.ac.in

30	NiCrFeTiAlTaNb	[18]
31	NiCCrCoMoWTaTiAlBHf	[19]
32	NiCoCrMoHfTaAlWReCZrB	[20]
33	NiCoCrAlTiZrNbMoWTaHfBC	[21]
34	NiCoCrMoFeSiMnCAlTiCuPBSZr	[10]

Structural features obtained using Image processing

The Table S2 below describes the geometrical features which are evaluated using various image processing filters and algorithms. These features are considered as structural descriptors.

Table S2: Structural descriptors

Sl.	Features	Description	Notation used
1	Volume (area) fraction	The ratio of number of areal pixels in white section (γ' phase) to the total number of pixels ($\gamma + \gamma'$ phase) in the binarised image.	VF
2	Boundary fraction	The ratio of number of boundary pixels in the white section (γ' phase) to the total number of pixels ($\gamma + \gamma'$ phase) in the canny image.	BF
3	Effective area	The number of pixels contained in white section (γ' phase) calculated over the number density of γ' phase.	A_{eff}
4	Effective perimeter	The number of pixels on the border of white section (γ' phase) calculated over the number density of γ' phase.	P_{eff}
5	Effective radius	The average radius of white section (γ' phase) calculated over the number density of γ' phase.	R_{eff}
6	Shape factor or circularity	The compactness of the area in the white section (γ' phase) calculated over the number density of γ' phase.	SF
7	Aspect ratio	The ratio of minor to major axis of the minimum enclosing rectangle that surrounds the white section.	AR
8	Density of grains	The number of grains (γ' phase) included in the image.	N_{obj}

List of elemental features

The table below illustrates the composition-weighted averaged elemental features, which are considered as descriptors.

Table S3: Elemental descriptors

Sl.	Elemental features	Notation used
1	Melting point	T_M^{mean}
2	Boiling point	T_B^{mean}
3	Specific heat	C_V^{mean}
4	Pauli electronegativity	E_p^{mean}
5	Electron affinity	E_a^{mean}
6	Valency	V^{mean}
7	First ionisation energy	E_1^{mean}
8	Covalent radius	r_{cov}^{mean}
9	Density	ρ^{mean}
10	Bulk Modulus	B^{mean}
11	Thermal Conductivity	κ^{mean}

Standardization

Figure S1 shows the standardization of features by subtracting the mean and dividing by the standard deviation for the training data.

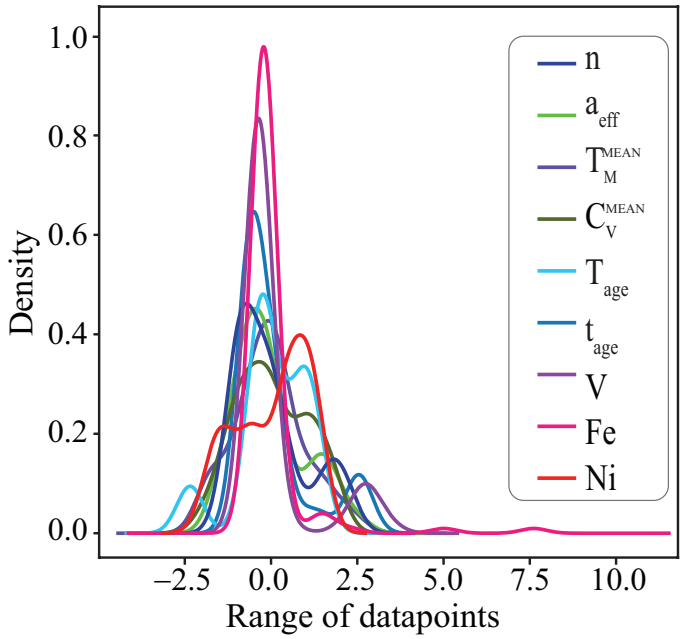

Fig. S1 Standardization of data for preprocessing of ML model

Machine learning models

Fig S2 shows the performance of different ML models on the best feature set (feature set-III). The GPR model outperforms giving the highest R^2 of 0.99/0.98 and lower rmse of 0.13/0.15 for train/test data.

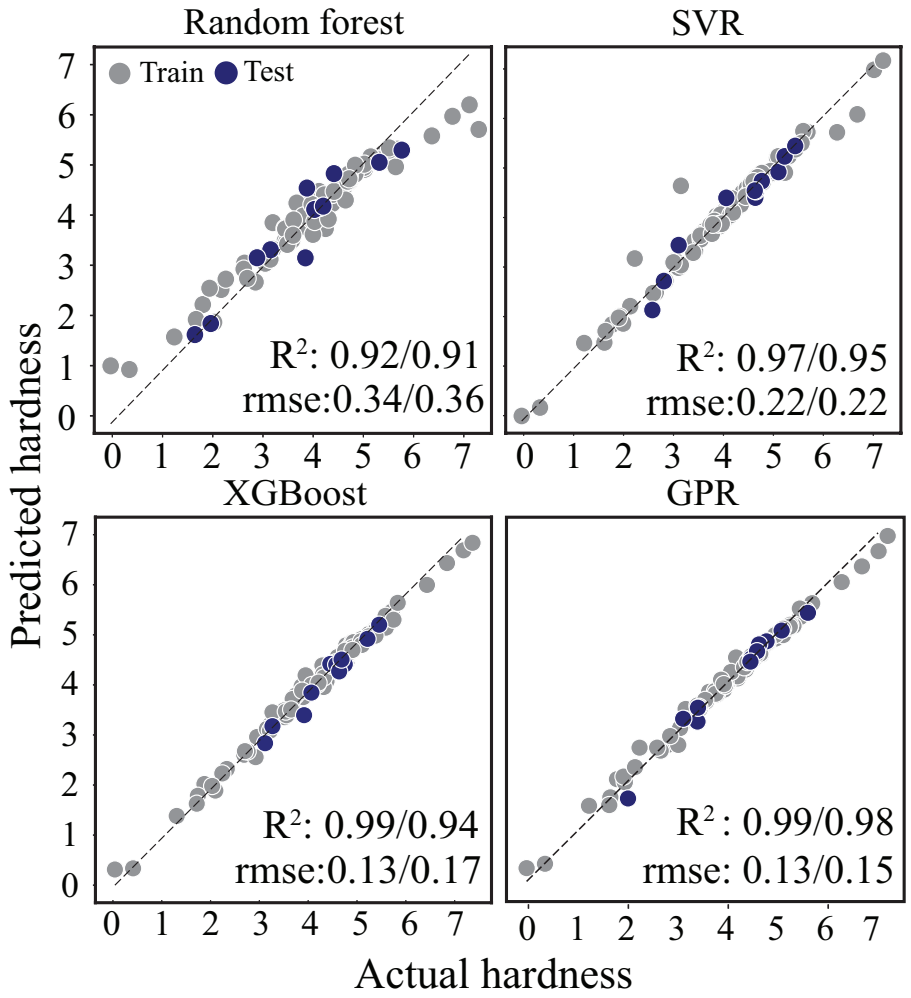

Fig. S2 Performance of different models on feature set-III

Learning Curve

Figure S2 shows the learning curve for the developed model with feature set-III. As a function of increasing training data, the train and test R^2 and rmse converge around 0.98 and 0.15, respectively, showing no overfitting for 90 percent of training data.

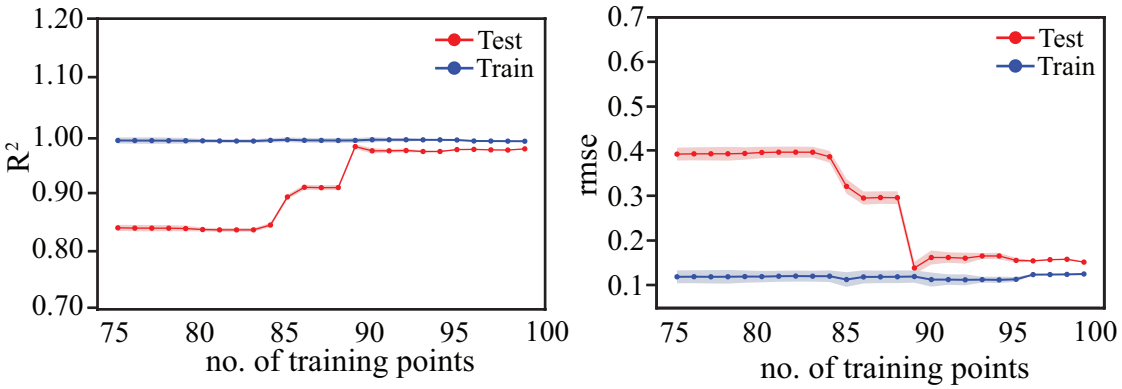

Fig. S3 Learning curve for different training datapoints for Feature set-III

Validation of ML model

For the validation of our best model, some random SEM microstructures of Co- and Ni-based superalloys [22, 23,24] are collected and the best GPR model is employed to predict Vickers hardness. The table below shows the predictive capability of our model on unseen data. Further, from the residual plot shown in Fig S4, it can be ascertained that the prediction error is minimal. Thus, our developed model can be utilised to predict Vickers hardness for any unseen data.

Table S4: Validation of GPR model

Sl.	Actual (Experimental) hardness	Predicted (ML) hardness
1	3.84	5.05
2	4.59	5.17
3	4.16	2.92
4	4.13	4.09
5	3.91	3.46
6	3.92	3.14
7	3.97	2.79
8	3.61	2.88

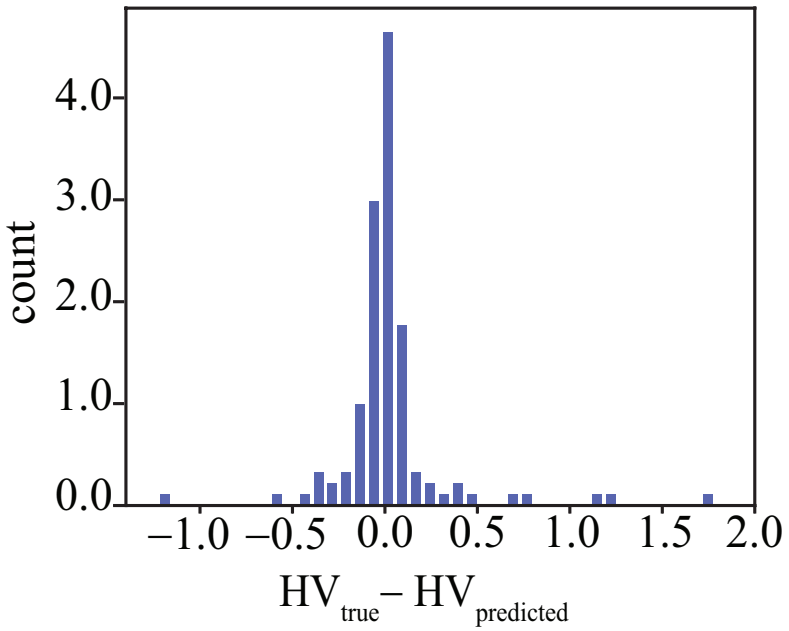

Fig. S4 Residual plot for validation of unseen data for the prediction of Vickers hardness.

References

1. P.M. Mignanelli, N.G. Jones, M. Hardy, H.J. Stone, Materials Science and Engineering: A **612**, 179 (2014)
2. P.J. Bocchini, C.K. Sudbrack, D.J. Sauza, R.D. Noebe, D.N. Seidman, D.C. Dunand, Materials Science and Engineering: A **700**, 481 (2017)
3. J. Sato, T. Omori, K. Oikawa, I. Ohnuma, R. Kainuma, K. Ishida, Science **312**(5770), 90 (2006)
4. H.Y. Yan, V. Vorontsov, D. Dye, Intermetallics **48**, 44 (2014)
5. Y. Nunomura, Y. Kaneno, H. Tsuda, T. Takasugi, Acta materialia **54**(3), 851 (2006)
6. K. Kawahara, Y. Kaneno, A. Kakitsuji, T. Takasugi, Intermetallics **17**(11), 938 (2009)
7. C. Barbosa, J. Nascimento, I. Caminha, I. Abud, Engineering Failure Analysis **12**(3), 348 (2005)
8. T. Osada, N. Nagashima, Y. Gu, Y. Yuan, T. Yokokawa, H. Harada, Scripta Materialia **64**(9), 892 (2011)
9. R. Mitchell, M. Preuss, S. Tin, M. Hardy, Materials Science and Engineering: A **473**(1-2), 158 (2008)
10. M. Göken, M. Kempf, Acta Materialia **47**(3), 1043 (1999)
11. Z. Zhong, Y. Gu, Y. Yuan, Z. Shi, Materials Letters **109**, 38 (2013)
12. T. Osada, Y. Gu, N. Nagashima, Y. Yuan, T. Yokokawa, H. Harada, Acta Materialia **61**(5), 1820 (2013)
13. G.A. Zickler, R. Schnitzer, R. Radis, R. Hochfellner, R. Schweins, M. Stockinger, H. Leitner, Materials Science and Engineering: A **523**(1-2), 295 (2009)
14. C. Cui, D. Ping, Y. Gu, H. Harada, Materials transactions **47**(8), 2099 (2006)
15. X. Li, J. Bai, P. Liu, Y. Zhu, X. Xie, Q. Zhan, Journal of Alloys and Compounds **559**, 81 (2013)
16. H. Li, X. Song, Y. Wang, G. Chen, Rare Metals **29**(2), 204 (2010)
17. Y. Zhang, H. Fu, X. Zhou, Y. Zhang, J. Xie, Materials Science and Engineering: A **737**, 265 (2018)
18. A. Picasso, A. Somoza, A. Tolley, Journal of Alloys and Compounds **479**(1-2), 129 (2009)
19. H. Peng, Y. Shi, S. Gong, H. Guo, B. Chen, Materials & Design **159**, 155 (2018)
20. L. Murr, E. Martinez, X. Pan, S. Gaytan, J. Castro, C. Terrazas, F. Medina, R. Wicker, D. Abbott, Acta Materialia **61**(11), 4289 (2013)
21. H. Kim, S. Chun, X. Yao, Y. Fang, J. Choi, Journal of Materials Science **32**(18), 4917 (1997)
22. H. Wu, X. Zhuang, Y. Nie, Y. Li, L. Jiang, Materials Science and Engineering: A **754**, 29 (2019)
23. C. Che, S. Yang, M. Wei, L. Zhang, Q. Li, J. Gao, Y. Du, Journal of Mining and Metallurgy, Section B: Metallurgy **53**(3), 303 (2017)
24. F.L.R. Tirado, S. Taylor, D.C. Dunand, Acta Materialia **172**, 44 (2019)