Advances in Additive Manufacturing of Materials: Current status and emerging opportunities

Prof. Bikramjit Basu

Materials Research Center, Indian Institute of Science, Bangalore

Lec48

We will continue our discussion on the emerging topics in 3D printing or additive manufacturing and that is data science. what we have learnt in the last few lectures essentially what are the overview, what are the definitions of this machine learning and deep learning. And then how to define data sets, what are the scientific ethical things then one has to follow in this artificial intelligence in terms of the biases and it should be done everything is an unbiased manner and so on and so forth,   those things are very important for us to recall. Now overview of the artificial intelligence, so this is essentially as I defined in one of the last lecture that is a scientific discipline that uses algorithms and techniques leveraging computer to mimic human intelligence. here this has been mentioned earlier to mimic human intelligence. Now what is machine learning? Algorithms whose performance improves over time and as they are exposed to more and more data.

More and more data means quantum of data sets. Improves over time means you keep on adding more and more high quality data sets in your training data sets and then you can learn the model better. Then there are supervised learning, unsupervised learning and there is a reinforcement learning. And fourth one is the deep learning.

deep learning essentially is a subset of machine learning and where multilayer neural networks learn from a vast amount of data. deep learning essentially uses artificial neural networks and this is multiple layers. Now, if you look at this biological structure of the neurons, so you have a axons, you have a long axons, then you have a cell body and this is your cell body and often they call and then you have your dendrite and this is your nucleus, and this is your axons. and the tail part you have a synapses. these particular neurons is actually is kind of a biological symbols which are used to explain that what is the meaning of neural networks like the way you can see that synapses and dendrites they kind of network with the other cells in their cellular microenvironment.

Similar way multilayer neural networks essentially learn from a vast amount of data. I have mentioned it earlier I am mentioning it again. deep learning and machine learning are not synonymous and in case of deep learning you certainly require much more larger quantum of data than you need for the machine learning. One of the things which we learnt

also that AI/ML in which cases it becomes very important. It is most relevant in absence of any established relationship between dependent and independent variables.

If there is no established relationship, for example, you know, force is equal to mass times acceleration, this you learn from Newton's second law. there is a relationship between force and acceleration and this relationship what we learn from a basic physics is that force applied is directly proportional to change of momentum. v-u by t and if it is a momentum so then you write down that F is equal to Kma and then you can consider given certain assumption if you get K is equal to 1 then it comes out to be mass times acceleration. But there is a direct proportional relationship here but in the case of 3D printing and any other cases when there is no established relationship how the defect that will be generated in the parts will be directly related to laser beam power or scan speed and so on or what would be the laser scan speed or beam power that will generate this is the depth of this melt track you know there is no direct relationship . therefore in those kind of cases AI/ML actually plays an important role.

Researchers have used AI/ML for multiple different classes of materials and one of the classes of materials they have used for multi component glasses. glasses are normally silica based amorphous solids. you can have silica, you can have multiple oxides like silica, alumina, calcium oxide, phosphorus pentoxide, so on, what you see here in this simple glass composition I am seeing that is a four different oxides, amount of x, what would be the value of x, y, z, u, v, w of different oxides that actually requires intuitive tailoring, that actually requires significant amount of experiments to be conducted this is one of the problem that you know one of my colleague from Pennsylvania United States John Mauro and his group and collaborators that they have been working for quite some time and they have essentially come up with the field called glassomics. glass and omics so they have put it together and then they have established this particular domain or theme research theme. in the glasses they are used at very different applications wide spectrum of applications.

Glasses based windows are used in the space exploration, space applications, they are used in mobile, they are also available in different colors and so on. And what you can see that these glasses by definition they are amorphous solids. essentially they do not have any long range order but they have a short range order. And their strength property is important because they are essentially very brittle that means they are prone to fracture very easily and that is why they have a very small K1c values that is a fracture toughness much lower than traditional or conventional ceramics. classes are multi-component oxides, so we need to know what would be the different values of X, Y, Z in terms of the different oxide addition alumina, B_2O_3, Na_2O, MgO, CO, ZnO and this combination of oxides that will give rise to some kind of compressive stresses in the as manufactured classes.

Now, when the glass is being fabricated or manufactured, there are different factors that are important. One of the factors is liquidus viscosity or zircon breakdown viscosity. Another factors can be ion exchange time. these are the parameters that essentially influence the quality of the glass materials that you are developing. machine learning has been used very successively or at least Mauro and their group have pioneered the machine learning in the field of multicomponent glasses where there are different product attributes which I have mentioned hardness, Young's modulus, coefficient of thermal expansion, density these are like material parameters and different manufacturing attributes like liquidus viscosity, forming temperature, refractory compatibility, melt risk rate, etc were also   used.

 In the next few slides, what I am going to do, I am going to essentially show you scientific case study on machine learning for melt pool predictions and this is the case for the DED directed energy deposition of stainless steel 316L. directed energy deposition, this has been shown to you before but I think many lectures before, so I thought that it will be good idea to show you this video so that it will remind you in this context of the machine learning that how this laser beam power that can be changed and there you know that your powders are also getting synchronized in a very high throughput manner and in the top you can see this laser beam also is kind of stationary but this support is actually rotating, right. in this particular case what we are trying to see the design and toolpath synchronization is important and because without synchronization you cannot get the good high quality product at the end of the DED process, what are the parameters that are important on material related? These are like spherical powders, And the spherical powders have a specific D10 value, D50 value, D90 value. D50 value is like 60.

54 or 60.5 micron and you have a gamma austenite. what we are trying to do suppose you vary the different combination of parameters in stainless steel and then you get this different large number of single track stainless steel 316 wall track and each track has a combination of values of h is the weld pool height, d is the weld pool depth, w is the weld pool width, ai is AL is the lower weld pool area and AU is the upper weld pool area. this is the selected parameters that we have used in this 400 to 1000 watt, scan speed varies from 600 to 1400 and powder feed rate is 2,3,4,5. And this particular DOE essentially will generate large number of single track experiments and single tracks of SS316. And then SS316 this particular different color dotted area essentially denotes this is the some combination for example one particular laser power, one particular laser speed and different   combination   of   powder   fitted   for   example.

 we can get a certain combination of this one. And then we have measured the different morphological parameters in terms of D value, L value, D value, W value and so on and

this shows some of the representative cross section for the printable and defective region. when you consider the geometrical parameters in terms of the width, height and depth and you can see that this is the quantum of data sets that we have generated in terms of geometrical parameters. And when you look at the probability density, for example, height, it is very skewed kind of distribution. If you look at the depth, it is a much broader distribution.

If you look at that width, it is also equally broader distribution. Now, upper area, area upper and area lower, area upper is essentially much narrow skewed distribution, but area upper is also little broader distribution with a long tail. we can essentially build up this Pearson'smap here just to show different parameters power, scan speed, powder flow rate, width versus similar parameters here as has been shown. what are the important outcome of this analysis is the distribution of different geometric parameters were identified and Pearson's correlation coefficient that is the Pearson's map as I am saying, Pearson's correlation coefficient map that shows the relationship among the process parameters. this particular slide also again I have shown you before that we have used k cross fold validation that is k fold cross validation and here k value is 5 and we have this 96 data points.

We have a combination of laser power, scan speed and power of feed rate that gives you the combination of values, output is a width, depth, height and area. this is the set of data that we have generated using Gaussian process regression, this is one of the ML models, Gaussian process regression. what you see that you have a width, depth, height, area upper and area lower. in this particular case, if you see the predicted value and versus true value, testing and training. training value is much larger because training is your 80% datasets and testing is your 20% datasets.

what is the values you are getting in terms of performance matrix? R square is 0.97/ 0.97, that means for both the training and testing that R square value is 0.97. RMSE is very, very small.

It is 0.001. this kind of combination is quite acceptable. But one has to also show that what is the residual error distribution function distribution and here a residual error is distributed this is that blue is the test and then other one is the train and then you can see their maximum probability distribution is centered at origin 0.0 and it shows almost like a Gaussian distribution. Now when you go for the depth one again in the training and testing you get this R square and RMSE values when you have used GPR algorithm you are getting 0.

96, 0.96 it is fairly good values and RMSE is equally low is 0.003. If you go for the height one so again as you know that is the depth is the d. width is the W of the laser track width

and   then   if   you   have   this,   this   is   the   height.

 laser weld pool height is h. in the height case again the r square is 0.85 and 0.78 height is not that greatly correlated or this model does not perform as well as like for the example of width and depth. Upper area and lower area you can see that R square value is 0.

97/ 0.97, RMSE value is 0.002. and this is the area apart this is 0.94 and 0.87. in all the cases if you can see the residual error distribution is quite symmetric. it is of the Gaussian nature   and   it   is   almost   centered   around   0.

0. Therefore, what we concluded that Gaussian process regression is one of those successful machine learning algorithms which we can perhaps use for this particular directed energy deposition when we have essentially used 96 less than even 100 data points or data sets. less than 100 data sets, GPR improves, GPR essentially shows very good performance. Remember as I said that machine learning performance depends on time or improves over time when this machine learning models are exposed to more datasets. what it means that if the DED experiments is conducted even to generate larger datasets like more than 100 to 150 or 200 datasets. Then maybe some other ML algorithm performs even   better   than   that   of   the   GPR.

 that is the message I am trying to send across while concluding this particular slide. Now this essentially shows that what are the other algorithms which are used like LR linear regression, GPR is the Gaussian process regression, GBM that is gradient boosting machine. Now if you look at that five different parameters let us say for example width, R square and RMSE combination, GPR essentially it gives you as good as gradient boosting machine that is GBM. If you look at the depth, depth again these two GPR and GBM is better, if you good at Compared to linear regression, other two are still even better. If you do   this   area   track,   again   GPR   and   GBM   is   better.

 If you go to the melt pool, area melt pool, again GPR and GBM is better. GPR/ GBM, both perform equally well. in modeling this kind of different single track data in the directed energy deposition process. this is a kind of summary. summary essentially what you see here is that you know that you have this input data that is the width ,depth, height, upper   area   and   lower   area.

 You have the laser power, you have the scan speed, you have the powder feed rate, less than 100 data points and then you get these predictions and then in the predictions you get the training data sets and training data sets and testing data sets is 20%. And then you can see that what is the DED parameters, you can essentially predict what is the unknown process parameters and to see you would be able to see that what is the statistically reliable

prediction of this w,d,h,A_u,A_l. the green one is your observations, red one is your essentially predictions and this one is your uncertainty. this is the uncertainty region. this hatched region is your uncertainty region.

This red one is your predictions like testing datasets and this one is your training datasets. What we learnt from here is this machine learning at least three different algorithms what I have shown you, linear regression, Gaussian process regression and gradient boosting machines. In all these 3 cases what we have learnt that both the Gaussian process regression and GBM given that less than 100 data sets, so to be precise it is like 96 data sets. They are able to perform quite well and then you cannot distinguish in terms of the performance between GPR and GBM in terms of R square RMSE. And in the last slide, you have even seen that in all these cases, your residual error analysis, your residual error distribution also shows Gaussian nature.

And it shows the maximum probability density is centered around 0.0, that is the origin. essentially I will come back to the summary slide. This summary slide essentially shows that you have the input data and this input data based on the different combination of this laser power, scan speed, powder feed rate and so on. These are called features like features used in manufacturing and these are the input data w,d,h,A_u,A_l.

you have the predictions versus features and these predictions versus features are essentially these features are like different features that I have mentioned here . if you put these values together then what we get these are your observations and this is your predictions this red ones. And this Gaussian distribution is that essentially residual error distributions either around the predictions or around the observations. And this residual error distribution is shown in the Gaussian is of the Gaussian type . as I said before that your 80% is training data set, your 20% is the testing data set.

Testing data set is also used for the validation of your training outcome. And you have, for example, you have a combination of unknown process parameters. And you want to predict that what would be the combination of unknown parameters to get a particular value of weld pool depth or height or width, that is possible using this particular GPR model Gaussian process regression. Similarly, we can also use gradient boosting machine is another alternative algorithms which performs equally well like Gaussian process regression. in the next lecture, I am going to essentially start with another scientific case study to substantiate our discussion on this emerging topic like artificial intelligence machine learning in the context of additive manufacturing. Thank you.