Artificial intelligence (AI) and machine learning (ML) techniques have been applied to support certain tasks in material science research. Some of those are: accelerated simulation, to predict properties of new materials, planning the route to synthesize materials and experimental parameter optimization (Mueller 2016, Butler 2018, Sha 2020). In the framework of this paper, we are interested in experimental design algorithms which are aimed to reduce the number of required experiments (Dieb, 2018). The approach implies generating a set of candidates (experiments), each candidate is evaluated according to an objective function and the most promising candidates are selected to be performed at the laboratory. In that way, the idea is to conduct only promising experiments, saving resources. An important element is how to assess the candidate experiments. To do so, we propose to use the data of some experiments already performed to train a model to predict a valuable property, in this case the absorbance. Then, through the model, evaluate new candidate experiments to decide which ones are interesting to be carried out.

We will use the *supervised learnin*g approach. With it, we use labeled data, to infer a model. The labeled data set is used as the basis for predicting the values of other unlabeled data through the use of machine learning algorithms. There are several supervised learning methods. In this work, we tested Support Vector Machine (SVM), Linear Regression (LR), and Artificial Neural Networks. The best result was given by the ANN, thus we concentrate here on the foundations of it.

In few words, an ANN is a network of processing units called artificial neurons. Neurons are connected following a certain topology. Fig. XXX illustrates a typical Multi-Layer Perceptron (MLP). Each neuron performs two operations. The first one is computing the *net input function*, which is a weighted sum of all the inputs to the neuron. If **x**={x1, …, xn} are the input values and **w**={w1, …, wn} are the weights, S=SUM1 n (xi\*wi). The result is input to an activation function f(S), which will generate an output of the neuron. It is common to use a logistic regression function f(s)=1/(1+e^-s), although other functions are also possible (ReLu, identity and tanh). The network is trained following what is called a backpropagation algorithm, which implies iterating until a certain error value is obtained.

The dataset used to train the model is comprised of 632 samples, each with 4 attributes or features: treatment time, atomic concentration of the nanoparticle and type of bacteria. In each case we have a measurement of the absorbance. The Orange ML platform (Demsar 2013) was used to generate several models and compare them. Four activation functions were selected, and three solvers, five alpha values and eight max iteration values. The performance of the models was evaluated using root mean square (RMSE), mean absolute error (MAE), mean square error (MSE) and coefficient of determination (R2). Table 1 shows the results for *S. aurea*, and Table 2 shows the results for *E. Coli*. The best setup is marked in bold. For *S. aureus* the best is a neural network with 121 hidden neurons, ReLu activation and L-BFGS-B solver. R2 = 0.983, which indicated a good representation of the data. The error is approximately 2%. In the case of *E. Coli*, the best result is a neural network with 140 hidden neurons, trained with ReLu activation and L-BFGS-B solver. The results show R2= 0.9870, with an error of 2.3%.

Table 1, *S. aureus*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number Hidden Neurons | Activation | Solver | Alpha | Max iterations | RMSE | MAE | R2 |
| 200 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0310 | 0.0210 | 0.9820 |
| 250 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0300 | 0.0200 | 0.9830 |
| 150 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0300 | 0.0200 | 0.9830 |
| 121 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0310 | 0.0200 | 0.9830 |
| 121 | Identity | L-BFGS-B | 0.0001 | 100 | 0.1880 | 0.1433 | 0.0339 |
| 121 | Logistic | L-BFGS-B | 0.0001 | 100 | 0.1880 | 0.1430 | 0.3390 |
| 121 | Tanh | L-BFGS-B | 0.0001 | 100 | 0.0370 | 0.0280 | 0.9750 |
| 121 | ReLu | SGD | 0.0001 | 100 | 0.2070 | 0.1470 | 0.1980 |
| 121 | ReLu | Adam | 0.0001 | 100 | 0.1180 | 0.0790 | 0.7390 |
| 121 | ReLu | L-BFGS-B | 0.0005 | 100 | 0.0310 | 0.0200 | 0.9820 |
| 121 | ReLu | L-BFGS-B | 0.01 | 100 | 0.0310 | 0.0200 | 0.9820 |
| 121 | ReLu | L-BFGS-B | 0.1 | 100 | 0.0310 | 0.0220 | 0.9820 |
| 121 | ReLu | L-BFGS-B | 1 | 100 | 0.0660 | 0.0430 | 0.9190 |
| 121 | ReLu | L-BFGS-B | 0.1 | 70 | 0.0320 | 0.0240 | 0.9800 |
| 121 | ReLu | L-BFGS-B | 0.1 | 85 | 0.0320 | 0.0230 | 0.9810 |
| **121** | **ReLu** | **L-BFGS-B** | **0.0001** | **95** | **0.0300** | **0.0200** | **0.9830** |

Table 2, E. coli.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Number Hidden Neurons | Activation | Solver | Alpha | Max iterations | RMSE | MAE | R2 |
| 150 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0380 | 0.0270 | 0.9830 |
| 100 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0390 | 0.0280 | 0.9820 |
| 75 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0400 | 0.0290 | 0.9810 |
| 50 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0500 | 0.0360 | 0.9700 |
| 175 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0390 | 0.0280 | 0.9820 |
| 200 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0370 | 0.0270 | 0.9840 |
| 500 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0360 | 0.0260 | 0.9850 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0390 | 0.0280 | 0.9830 |
| 140 | Identity | L-BFGS-B | 0.0001 | 100 | 0.1680 | 0.1290 | 0.6680 |
| 140 | Logistic | L-BFGS-B | 0.0001 | 100 | 0.1680 | 0.1280 | 0.6690 |
| 140 | tanh | L-BFGS-B | 0.0001 | 100 | 0.0700 | 0.0530 | 0.9420 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 100 | 0.0390 | 0.0280 | 0.9850 |
| 140 | ReLu | SGD | 0.0001 | 100 | 0.1720 | 0.1310 | 0.6520 |
| 140 | ReLu | Adam | 0.0001 | 100 | 0.1110 | 0.0830 | 0.8540 |
| 140 | ReLu | L-BFGS-B | 0.0005 | 100 | 0.0400 | 0.0280 | 0.9820 |
| 140 | ReLu | L-BFGS-B | 0.001 | 100 | 0.0400 | 0.0280 | 0.9820 |
| 140 | ReLu | L-BFGS-B | 0.01 | 100 | 0.0400 | 0.0290 | 0.9810 |
| 140 | ReLu | L-BFGS-B | 1 | 100 | 0.0800 | 0.0580 | 0.9240 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 150 | 0.0340 | 0.0240 | 0.9860 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 200 | 0.0330 | 0.0220 | 0.9870 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 250 | 0.0330 | 0.0220 | 0.9870 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 500 | 0.0330 | 0.0210 | 0.9870 |
| 140 | ReLu | L-BFGS-B | 0.0001 | 75 | 0.0450 | 0.0330 | 0.9760 |
| **140** | **ReLu** | **L-BFGS-B** | **0.0001** | **165** | **0.0340** | **0.0230** | **0.9870** |

.

In a previous work (Navarro 2021), we have done this for ZnO dopped with Yterbium. In future work we would like to determine whether such methodology could be generalized to represent the effect of more rare earth elements.

References

Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) [Orange: Data Mining Toolbox in Python](http://jmlr.org/papers/volume14/demsar13a/demsar13a.pdf), Journal of Machine Learning Research 14(Aug): 2349−2353.

Mueller, T., Kusne, A.G. and Ramprasad, R. (2016). Machine Learning in Materials Science. In Reviews in Computational Chemistry (eds A.L. Parrill and K.B. Lipkowitz). <https://doi.org/10.1002/9781119148739.ch4>

Butler, K.T., Davies, D.W., Cartwright, H. *et al.* Machine learning for molecular and materials science. *Nature* **559,**547–555 (2018). <https://doi.org/10.1038/s41586-018-0337-2>

Dieb T.M., Tsuda K. (2018) Machine Learning-Based Experimental Design in Materials Science. In: Tanaka I. (eds) Nanoinformatics. Springer, Singapore. https://doi.org/10.1007/978-981-10-7617-6\_4

Sha, W., Guo, Y., Yuan, Q., Tang, S., Zhang, X., Lu, S., Guo, X., Cao, Y. and Cheng, S. (2020), Artificial Intelligence to Power the Future of Materials Science and Engineering. Adv. Intell. Syst., 2: 1900143. <https://doi.org/10.1002/aisy.201900143>

D. E. Navarro-López, R. Garcia-Varela, O. Ceballos-Sanchez, A. Sanchez-Martinez, G. Sanchez-Ante, K. Corona-Romero, D.A. Buentello-Montoya, A. Elías-Zuñiga, E. R. López-Mena, Effective antimicrobial activity of ZnO and Yb-doped ZnO nanoparticles against Staphylococcus aureus and Escherichia coli, Materials Science and Engineering: C, Volume 123, 2021, 112004, ISSN 0928-4931,https://doi.org/10.1016/j.msec.2021.112004.