Tutorial Question & Answers

1. Is there an intuitive reason for why FFNNs cannot capture resonance phenomena?
   **Answer:** There is no specific reason. However, you may need more samples around the resonance frequency to be able to capture it accurately which makes it more difficult. As was shown in the example (Problem 4), the error control you get saturates after a while and it is hard to get good accuracy.
2. When showing about predicting VCO voltage based on control voltage, there were two functions mentioned. Can both not be represented with one NN?
   **Answer:** That is possible however from a practical stand point it becomes difficult to implement for a VCO. Notice that there is a mod 1 operator being used on the second NN in Problem 2 which requires the two networks to be separated. For further details please see: *Huan Yu, Hemanth Chalamalasetty and Madhavan Swaminathan, "Modeling of Voltage-Controlled Oscillators Including I/O Behavior Using Augmented Neural Networks", IEEE Access, Volume: 7, pp: 38973 – 38982, 2019.*
3. For modeling IO circuits, what's the pros and cons of NN vs. IBIS model?
   **Answer:** Both NN and IBIS protect IP. However, NN provides for a better model fidelity since the waveform behavior can be captured directly from a transistor circuit unlike a simplified table as in IBIS.
4. Approximately, how many training samples are necessary to train a S-TCNN, for instance for 10 input parameters?
   Answer: Hard to say since it depends on the data samples and response. For further details please see: *Hakki Mert Torun, Huan Yu, Nihar Dasari, Venkata Chaitanya Krishna Chekuri, Arvind Singh , Jinwoo Kim, Sung Kyu Lim, Saibal Mukhopadhyay and Madhavan Swaminathan, "A Spectral Convolutional Net for Co-Optimization of Integrated Voltage Regulators and Embedded Inductors", IEEE/ACM International Conference on Computer-Aided Design (ICCAD), 2019*.
5. How do you tackle process variation (PV) phenomena during using ML algorithms for prediction when you dont have PV information, so you cannot train your network by applying PV information during training?
   **Answer:** ML is only as good as the data used to train it. I am assuming that PVT information in some form is available. For example, the process space can be used as input for training where the space covers the variations. If so, the inferences derived from the ML model will cover the process variations.
6. How big are the typical models e.g. number of layers, parameters?
   **Answer:** In the semiconductor and packaging space the number of parameters can be many (5-100), number of layers can be many (>20) etc. A lot depends on the complexity of the data and the patterns you are trying to derive from it which determines the breadth and depth of the NN.
7. When showing about predicting VCO voltage based on control voltage, there were two functions mentioned. Can both not be represented with one NN?
   **Answer:** Answered earlier.
8. How to find proper settings for the neural network? Like how many layers, neurons I should have?
   **Answer:** There is no easy way. It has come based on some initial work by following the flow chart I showed you on the slide titled "Steps Required for Building a NN"
9. Currently we are talking about small circuits and their frequency responses. How do such models pan out with respect to complex VLSI circuits such as microprocessors or microcontrollers?

**Answer:** The examples I showed you are real world examples which has been used to design hardware. These same concepts can be expanded to apply to for example the physical routing of large chips etc.

10. Though, there are various optimization algorithms available to us, but if problem is complex or taking high computational time, so can we solve the optimization problem using Artificial neural network?
    **Answer:** When you have a non-convex response surface ML based methods become very useful. Addressing computational time is an added benefit.
11. What kind of network structure do you think is good when 2D data such as the shape of the PDN inputs to the network and frequency data such as s-parameters outputs?
    **Answer:** Try the S-TCNN described in the presentation. The details of this paper are provided in one of the earlier questions.
12. What courses would you recommend to a student interested in this technique?
    **Answer:** There are many online tutorials available. Most are based on more of a CS approach. Follow these with presentations that are more hardware centric where you can understand how the methods developed in the CS community can be applied. A lot of the algorithms can be downloaded for free or are available as function calls in Matlab and others. Experiment with them.
13. How are the candidate points actually determined in the progression of TSBO and approximately how many candidate points are required to predict the next sample accurately?
    **Answer:** Candidate points are nothing but points selected based on a sampling strategy. There are many such sampling strategies available – one example is LHS. The other is the tree structure that was shown during the tutorial. The number of candidate points is a function of how many you use to represent a volume. These candidate points are used in the acquisition function to determine the next sampling point. This is an approach to avoid optimization of the acquisition function itself. Please see one of sampling strategies described here: *Madhavan Swaminathan, Hakki Mert Torun, Huan Yu, Jose Ale Hejase and Wiren Dale Becker, "Demystifying Machine Learning for Signal and Power Integrity Problems in Packaging", IEEE Transactions on Components, Packaging and Manufacturing Technology, Volume: 10, Issue: 8, pp: 1276 – 1295, 2020.*
14. Is confidence bound is derived from infinity curve fitting? If so, then how to get those infinity curves?
    **Answer:** You make use of the Gaussian Process (GP) that mimics the behavior of using infinite curves.