Abstract
The world’s population is on the rise and in order to feed the world in 2050, food production will need to increase by 70%. Today, the agriculture industry works to optimize the amount of food gained from plants by breeding plants with the strongest, highest-yielding genetics. It is believed data-driven strategies can help industry breed better seeds, faster. As a result, it is crucial to construct powerful predictive models for phenotype prediction based on genotype and environment data (so-called G by E problem) and this may help scientists more accurately choose seeds that increase the productivity of the crops and help address the growing global food demand.
Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers. In this presentation, we focus on how to use deep neural networks to predict crop yield as phenotype based on genetic and environment data. Then, we train deep neural networks and test them on real word dataset, 2018 Syngenta Crop challenge, to demonstrate their performance.

About the Speaker
Saeed Khaki is a Ph.D. student of the Industrial Engineering department at Iowa State University. He received his master’s and bachelor’s degrees in Industrial Engineering at Amirkabir University of Technology. His research interests are Machine Learning, Optimization and Stochastic processes. Saeed has recently been selected as one of the five finalists for the 2018 Syngenta Crop Challenge in Analytics.