

International Conference on

## COMPUTATIONAL BIOLOGY AND BIOINFORMATICS

September 05-06, 2018 Tokyo, Japan

**An ensemble machine learning applied to DNA sequence recognition****Kosuke Imamura**

Eastern Washington University, USA

Machine learning is becoming increasingly important today as it has a wide range of applications from medical image analysis to financial market prediction. There are various machine learning techniques. Yet, the reliability is difficult to define and measure, whether genetic programming or neural networks, a machine learning tool to be deployed is often the one which performs the best on sample data or an ensemble of learners. Ensemble learning is a software version of N-Modular Redundancy in hardware (NMR). NMR improves reliability because the probability of failure is much lower than a single device system. However, NMR assumes independent failures, which may not exist ensemble learning. If the learners acquire similar knowledge from given data sets, there is little performance gain, no matter how they are combined. So, the question is how we could realize NMR assumption of independent failures in machine learning. It is rather simple; first create as many learners as possible. Second, find a combination of learners that exhibits a failure rate close to the expected statistical failure rate. Such ensemble will give reasonable assurance that each learner acquired the knowledge from data independently of other ensemble members. We call this method Probabilistically Optimal Ensemble (POE). Cluster computing environment is ideal for this purpose since POE is computationally intensive. We show significant accuracy improvement by a POE method in E. coli DNA promoter region classification. Also in progress is protein solubility prediction by heterogeneous learners.

**Biography**

Kosuke Imamura has pursued his PhD in Computer Science from University of Idaho and has done his Postdoctoral Research in Protein Solubility Prediction at Molecular Kinetics, Inc. He has extensive industrial experience from medical patient monitoring system to flight simulator. He is currently a Computer Science Professor at Eastern Washington University, researching high-speed neural computation on reconfigurable hardware.

kimamura@ewu.edu

**Notes:**