**Title: Agricultural and Aquatic Applications of Machine Learning and Big Data**

Put Your Photo Here

***Dr. Charles C Zhou***

CEO

Cascade Clean Energy, Inc.

USA

***Abstract*** *(within 300 words)*

Data sciences are rapidly changing the face of biotechnology, with the largest application being both land based and aquatic plants. Our topic covers three big data set categories.

(1) The Agriculture big data set covers crops from nature (land and water) to ingestion

(2) The Health big data set covers food from ingestion to excretion

(3) The Environmental big data set then covers the nutrients after excretion, and back into nature

That completes the whole cycle of big data of life sciences, where, metagenomics — the genomics of organism communities reshapes the landscape of microbiology. Transcending individual genes and genomes of an organism, metagenomics soil, water or any samples from the uncultured nature offers access to all of the genomes in a community, revealing the secrets of the ‘uncultured swarm world’ — the number of microbial species that cannot currently be isolated into pure culture due to technical or economical barriers. Metagenomics will expand our ability to discover and benefit from microbial capabilities, improve our understanding of microbial communities and probably lead to major advances in medicine, agriculture, energy production and bioremediation of soil and waterways. We built our data model from USPTO database, NCBI database, JGI (Joint Genomic Database) and KEGG database, as well as our own bio-database.

To screen for smart microbes, we applied a machine learning system – “computer assisted strain construction and development engineering” (CASCADE) to a collection of organisms and plants, and found correlations and predictive patterns between the organisms’ genetic information and their metabolic behavior. We compiled and gathered five types of genetic information from public databases and our private database. We also compiled metabolic pathways from public databases and investigated metabolic reconstructions for the organisms with only genome information. We defined a measurement called “average metabolic efficiency” (AME) for a given organism and found it highly correlates with the metabolic capabilities in real life. That provides information on how a plant (such as crops, weeds or algae) interact with microbes, microsponges, or chemicals.

We developed three applications using the selected smart microbes or enzymes, 1) Microbial Fuel Cell to generate electricity from rotten algae or water weeds and 2) Methane Bioreactor to get clean energy from aquatic wastes 3) Composting the leftover wastes as fertilizers.

***Biography*** *(within 300 words)*

Dr. Charles Zhou, Doctor of Science from MIT in 1991, CEO of Cascade Clean Energy, Inc., Co-founder of Quantum Intelligence, Inc. in USA. While working for Intel from 1995 to 1999, he initiated a data warehouse project and introduced the data mining and machine learning technology to the Intel Pentium production line, which received extensive improvement of yield and economy. He started to build a life science database for biotechnology while he was leading an IBM team on Intelligent Miner for data and for Text, which are the foundations of IBM’s Watson and Smart City program. After September 11 of 2001, He led the Quantum Intelligence team for extensive data sciences and machine learning projects in biotech, environment and agriculture. He developed quantum intelligence system, collaborative learning agents, web crawler, and Cascade for agriculture, biotech, environment and health applications. He was chairing a science and technology team for the MITCNC in silicon valley, the club initiated a clean energy forum in 2003, and now a worldwide clean tech open forum from 2005 to today.