Proposal-*Evaluating nuclear meltdown associated with both external and internal factors using statistical learning techniques*

With the increasing demands of high quality air and environmental protection by society, the use of nuclear energy is becoming more and more essential in today’s energy industry. While the technology of using the nuclear fission to produce heat and electricity is no more an impassable problem for many countries in this planet, the potential risk as well as accidents associated with nuclear industry is now becoming more and more significant and scientists and engineers are both so urgent to study on evaluating those risks and further to find efficient way to avoid them.

Among those nuclear accidents occurred in human’s history, nuclear meltdown is inevitable to be mentioned. Chernobyl disaster, Fukushima catastrophe, and Three Mile Island Nuclear incident had already shown us how horrible and how much trauma to both society and industry that a nuclear reactor meltdown incident can cause. A nuclear meltdown is a severe nuclear reactor accident that results in core damage from overheating. According to the official definition by the International Atomic Energy Agency (IAEA), the term *nuclear meltdown* refers to accidental melting of the core, leading to either complete or partial collapse of the core of a nuclear reactor and has been defined as the most hazardous accident associated with nuclear energy industry.

This study aims to use data analytics knowledge to find the relationship between the existence of nuclear meltdown and both internal and external factor of the nuclear reactor station. There are around 70 nuclear meltdown incidents in human’s history since the atomic era, and 30 of them are to be chosen as sample data. Besides that, another 30 nuclear reactors station which have not experienced any core meltdown incidents are also to be chosen as sample data. For external factors of the nuclear reactor station which can be seen as predictors, may include the topography of the land which the reactor was built, or the probability of the big natural disaster such as tsunami or earthquake that can happen to the site where the reactor was built. For external factors which can also be seen as predictors, may include the type of the nuclear power reactor, active core height, active core diameter, the duel inventory, the vessel type, the type of coolant, the type of moderator, the form of fuel, etc. These predictors all relate to the internal design of the reactor thus can be seen as internal traits of nuclear reactors.

By using statistical learning techniques, such as subset selection and stepwise model selection procedures to identify most significant factors that can contribute most to the happening of nuclear meltdown. What’s more, the logistic regression analysis can be used to train the model to predict the probability of the core meltdown for a future plant. The main target can be questioned as follow:

1. Are those factors all contribute significantly to the existence of nuclear meltdown?
2. Is there any possibility to predict the probability of a core meltdown incident based on those factors, both internal and external?
3. What is the likely relationship, or the best relationship between those values of factors and the probability of a core meltdown disaster?
4. How reliable is the model that might be used to predict the probability of a core meltdown disaster for future nuclear power plant?

These questions listed above are aimed to be answered in this study, and the results could potentially assist both industry-relative government agency official and nuclear engineering engineers with their decisions in choosing appropriate locations and designs for nuclear power plant in the future.