[EJ] Evening Poster | M (Multidisciplinary and Interdisciplinary) | M-AG Applied Geosciences

[M-AG32]Marine Earth Informatics

convener:Seiji Tsuboi(JAMSTEC, Center for Earth Information Science and Technology), Keiko Takahashi(Japan Agency for Marine and Earth Science and Technology), Masaki Kanao(国立極地研究所)
Wed. May 23, 2018 5:15 PM - 6:30 PM Poster Hall (International Exhibition Hall7, Makuhari Messe)
In advancing the research of marine Earth science, observation and computer simulation is an essential element. In recent years, the performance of the observation apparatus is dramatically improved, along with the means of observation is diversified. It is becoming possible to observe in a resolution, which was not imaginable so far. Such data to be generated from the observation is tremendously large in quantity and its quality is drastically improved. To handle these huge and high quality dataset for data analysis, we need to have a high speed and large memory computer system but such a system now becomes within reach in our hands by the recent dramatic improvement of high performance computer system. On the other hand, researchers who can use this kind of large-scale computer in their studies are still quite limited. In this session, we try to review the situation of observation data that has undergone a dramatic change regarded with both quality and quantity in recent years of marine Earth science research. We also try to review the situation from a professional standpoint of simulation about the status of the high performance computer system to analyze these 'big data'. Also we focus on the state of the art data analysis technique and aim to share the outlook from the professional standpoint of computational science and professional position of observation science about the future direction of the marine Earth informatics research.

[MAG32-P03]Earth Sciences big data analysis using Unsupervised Deep Learning, and challenge to "Earth Search".

*Daisuke Sugiyama[1], Masuo Nakano[1], Daisuke Matsuoka[1], Onishi Ryo[1], Tanaka Yusuke[1], Seiji Tsuboi[1]
(1.Japan Agency for Marine-Earth Science and Technology)
Keywords:Deep Learning , Data Mining, Earth Simulator (ES), Search Engine

Recent progress in high performance computer and measurement apparatus has produced huge amount of Earth science "big data", which almost exceeds our ability to process. For the analysis of these "big data", it was common to develop analysis technique for each case. This procedure has a merit in producing desired results but, on the other hand, it does not enable us to handle increasing data. Here we use Variational AutoEncoder, which is one of the model of artificial intelligence techniques, to perform unsupervised learning for large dataset of JAMSTEC's marine-earth science data. Unsupervised learning is a machine learning procedure, which does not provide answered questions and suggestions. Using this model for learning, we try to create an artificial intelligence, which may estimate inherited hidden signatures from cloud, atmospheric pressure, precipitation and ocean current data of JAMSTEC's NICAM model. This artificial intelligence may discern characteristics in data which we may not be able to find and may help us to discover new scientific knowledge, such as unknown laws or phenomena.
We also try to examine if neighborhood search of earth science data is possible or not by using search method of latent variables which the artificial intelligence has acquired through unsupervised learning. We estimate latent variables of past observation or simulated data in advance and keep them as a database of indexes. The we put an Earth status (observation data) of arbitrary space and time as an input and estimate latent variables to perform neighborhood search against these indexes. By using this

procedure we may create an artificial intelligence which enables us to search similar status as if we use Google search and we may call this as "Earth search".