

Research Trends and Opportunities in Machine Learning in Biotech & Health Sciences

Dr. Kolla Bhanu Prakash & Dr. B. Mahendran

Koneru Lakshmaiah Education Foundation

drkbp@kluniversity.in & drmahi_bt@kluniversity.in

Implementation of Machine Learning in every sector of Biotechnology is little scary, however efforts in implementing will open up new avenues in BIO-IT world. Major thrust areas in Biotechnology is Agricultural and Health sector. Agriculture mainly depends upon environmental conditions. Applying deep learning algorithms to predict early seasonal calamities in a form of pest attack, natural calamities can increase economic GDP in agricultural sector.

Agricultural income largely depends on agro-climatic conditions such as seed quality, soil quality, water availability and finally food preservation strategies. There are several public and private agencies who documented seasonal database, which are being implied to AI algorithms to educate farmers to choose a suitable crop. Machine learning algorithms in conjunction with Internet of Things (IoT) are being deployed not only to identify the pest attack, fungal infections but also helping to deliver insecticides and pesticides in a targeted manner using drone technology (<https://www.mindtree.com/sites/default/files/2018-04/Artificial%20Intelligence%20in%20Agriculture.pdf>). Along the lines AI based algorithm are being used for the identification and removal of weeds (Ahmed M. Tobal and Sahar A. Mokhtar, Journal of Computer Science 10 (8): 1355-1361, 2014). In spite of the modern technologies many of the farmers are committing suicide, due to the fact that the seeds and pesticides sold in the market are low yield and low efficient fertilizers and pesticides. Hence it is utmost important to design and develop a powerful artificial algorithm to avoid adulteration of seeds and other agro-economic products. Apart from seed quality, Machine learning algorithms need to be developed to identify the ripening time from the field to end consumer in natural conditions. Most of the time fruits are ripened using artificial gases or chemical to attract consumer. However, consumption of these food may damage neuronal growth of our children. Hence it is utmost import to develop an AI based algorithm to identify organic vs nonorganic ripening for the detection of qualitative and quantitative measurement of nutrient value, agricultural products for the health and well-being of humankind.

Biotechnological applications are generating large amount of data towards DNA analysis and Drug Discovery, by investing money in billion dollars. Both the methodologies are paving a way to clinical findings to treat human disorders. Drug discovery and disease associated mutation discovery is time and cost effective towards personalized medicine. Advancement in computational methodologies such as artificial intelligence and deep learning algorithms can be used in conjunction with bigdata analytical tools. The machine learning algorithms can be employed not only to identify genotypes vs wildtype, normal vs hybrid variety, sick vs healthy cell, grading diabetes, cancer but also to predict biological activities and toxicity predictions against drugs. These advancements in AI and ML are going to shift clinical genetics to Desktop diagnostics.

Some startups started their research based on Desktop Diagnostics based on ML algorithms. For example a startup named Benevolent Bio developed a ML algorithm to identify a drug candidate against motor neuron disease, which is verified by another startup named SITraN (Sheffield Transitional Institute of Neuroscience). Similarly, Switzerland based another startup Sophia Genetics is using AI based platform to diagnose many diseases based \$50-200 per test.

The following list of companies are using AI based platform for Biotech Applications

S.No	Name of the startup	Year of Establishment	Type of Disease
1	BenevolentAI	2018	Motor Neuron Disease
2	Sophia Geentics	2011	Oncology, hereditary Cancer, Metabolic Disorder
3	DNAlytics	2011	Rhematoid disease
4	Artery`s	2018	Medical Imaging to treat Heart Disease
5	Adynxl	2018	Post-surgery treatment against pain
6	Atomwise	2012	Ebola infection and Multiple sclerosis

Deep learning techniques can be employed to identify the organic vs non organic farming. Artificial ripening using chemical gasses can damage the public health by damaging the nervous system, which can be avoided by differentiating natural ripening and artificial ripening. Along the lines, many of the fruits are being coated with wax, which is difficult to

detect by natural eye. This can be accomplished by development of computer vision based on neural network algorithms can enhance threshold detection. Usage of natural and artificial colors in the food sector can be traced using AI models.

Deep learning can be used for design imaging and reading protocols specific to different organs, lesion types, and patient characteristics. With the growing demand of neuroimaging scanners in hospitals and institutes, radiologists roles and challenges are increasing. The manual interpretation suffers from inter- and intra-radiologist variance. In addition, emotion, fatigue, and other factors will influence the manual interpretation result.

After indepth literature review it has been observed that, Computer-Aided Medical Diagnosis are procedures in medicine to assist radiologists and doctors in the interpretation of medical images, which may come from CT, X-ray, ultrasound, thermography, MRI, PET, SPECT, etc. In practical situations, CAMD can help radiologists interpret a medical image within seconds. Conventional CAMD tools are built on top of handcrafted features. Recent progress on deep learning opens a new era that can automatically build features from the large amount of data. On the other hand, many important medical projects were launched during the last decade (Human brain project, Blue brain project, Brain Initiative, etc.) that provides massive data. Those emerging big medical data can support the use of deep learning.

It is especially important to develop deep networks to capture normal-appearing lesions, which may be neglected by human interpretation.

Some of the topics observed from latest literature are given below

- Robotic soft tissue surgery and Google Car
- Deep network architecture for CAMD and big medical data
- Deep learning for cancer location, cancer image segmentation, cancer tissue classification, cancer image retrieval
- CAMD for neurodegenerative diseases, neoplastic disease, cerebrovascular disease, and inflammatory disease

As mentioned in Nature Biotechnology [Sept 2018], DeepVariant uses convolutional neural networks to improve the accuracy of variant calling.

In another article from Nature Medicine, authors mentioned that Intracortical activity data recorded over 2 years in a tetraplegic patient is used to develop an artificial intelligence algorithm that achieves fast, accurate, and stable movement decoding to reenact real-time control of the paralyzed forearm. Nature Medicine, 1–8 [Sept 2018]

Another Research team observed that DeepSequence is an unsupervised deep latent-variable model that predicts the effects of mutations on the basis of evolutionary sequence information. Results published in Nature Methods 15, 816–822. [Sept 2018]

At last to conclude I am very much excited to see how deep imaging will accelerate to a level that it will re-invent the future of healthcare and modern engineering applications.

Few latest openings for Machine Learning Researchers in Biotech & Health Sciences

<https://www.indeed.com/viewjob?jk=da418c559db641c7&tk=1cugcisf0busm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=3781024a968cae02&tk=1cugcisf0busm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=a153e30a5c737ee0&tk=1cugcisf0busm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=f1f7e33c9cac8cd8&tk=1cugcisf0busm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=8800fe28344a8c91&tk=1cugcisf0busm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=dc9b34515128ec6e&tk=1cugcov1ibusm803&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=f0966ecf3e2427c5&tk=1cugcr6ejbusm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=8f497dc64896aee1&tk=1cugcr6ejbusm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=122cfa78fa813328&tk=1cugcr6ejbusm802&from=serp&vjs=3>

<https://www.indeed.com/viewjob?jk=12fcce400a9a9507&tk=1cugcr6ejbusm802&from=serp&vjs=3>

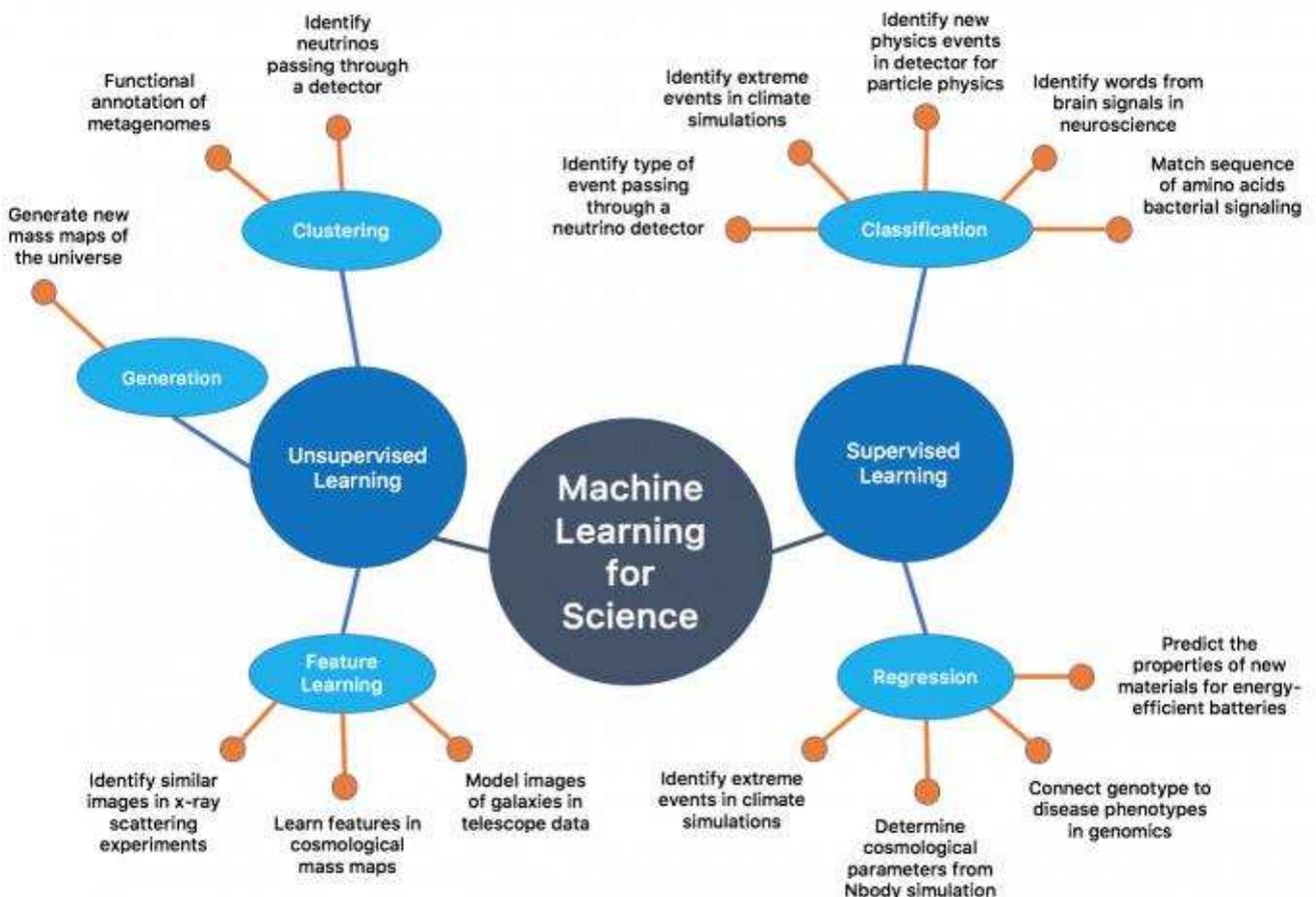
About the authors



Dr. Kolla Bhanu Prakash is working as Professor and Research Group Head for A.I and Data Sciences Research group in CSE Department, Koneru Lakshmaiah Education Foundation. Dr. Kolla Bhanu Prakash has 12+ years of experience working in the academia, research, teaching, academic administration. His current research interests include Machine learning, Deep Learning, Media Mining, Image Processing and Natural Language Processing. He was the recipient of Best Speaker award during M.Sc. He has authored over 4 books and 32 research papers in various national and international journals and conferences. His publications are indexed in Web of Science, Scopus, DBLP and Google scholar.



Dr. Mahendran Botlagunta, is an UGC-Research awardee from Koneru Lakshmaiah Education Foundation. His research focuses on interdisciplinary areas covering Cancer, Nanomedicine and Bioinformatics. His works were published in many peer reviewed international journals. His laboratory is currently engaged in detection of diseases using Machine Learning algorithms.



Source: <http://www.nersc.gov/users/data-analytics/data-analytics-2/deep-learning/>