Project Title:	Application of Artificial Intelligence in Disease Modelling
Project Number	IMURA0885

Reliance Industries Limited

Monash Main Supervisor (Name, Email Id, Phone)	Dr Sonika Tyagi, sonika.tyagi@monash.edu	*Full name, Email*
Monash Co-supervisor(s) (Name, Email Id, Phone)		
Monash Head of Dept/Centre (Name,Email)	Prof Terence O'Brien, Head of school	*Full name, email*
Monash Department:	Central Clinical School, Faculty of Medicine, Nursing and Health Sciences (FMNHS)	
Monash ADRT (Name,Email)		*Full name, email*
IITB Main Supervisor (Name, Email Id, Phone)	Dr Mithun Mitra, mithun@phy.iitb.ac.in	*Full name, Email*
IITB Co-supervisor(s) (Name, Email Id, Phone)		*Full name, Email*
IITB Head of Dept (Name, Email, Phone)	Prof. S. Umasankar, hod@phy.iitb	*Full name, email*
IITB Department:	Department of Physics	

Research Clusters: Research Themes:

Highlight which of the Academy's CLUSTERS this project will address? *(Please nominate JUST __one.__ For more information, see www.iitbmonash.org)*	Highlight which of the Academy's Theme(s) this project will address? *(Feel free to nominate more than one. For more information, see www.iitbmonash.org)*

1	*Material Science/Engineering (including Nano, Metallurgy)*		***1***	==*Advanced computational engineering, simulation and manufacture*==
2	*Energy, Green Chem, Chemistry, Catalysis, Reaction Eng*		2	*Infrastructure Engineering*
3	*Math, CFD, Modelling, Manufacturing*		3	*Clean Energy*
4	==*CSE, IT, Optimisation, Data, Sensors, Systems, Signal Processing, Control*==		4	*Water*
5	*Earth Sciences and Civil Engineering (Geo, Water, Climate)*		5	*Nanotechnology*
6	==*Bio, Stem Cells, Bio Chem, Pharma, Food*==		6	==*Biotechnology and Stem Cell Research*==
7	*Semi-Conductors, Optics, Photonics, Networks, Telecomm, Power Eng*		7	*Humanities and social sciences*
8	*HSS, Design, Management*		8	Design

The research problem

Define the problem

Healthcare research and development projects are utilising high throughput technologies to generate genome scale data at DNA, RNA, protein and metabolite level. Additionally, clinical and other digital health data is also abundantly available. Large-scale integration of these various types of biological and health data is required to design personalized medicine and therapies. Although accessible biomedical data is growing in quantity, heterogeneity and dimensionality, the methods to link information between data sources is lagging behind. The field of Artificial Intelligence (AI) and machine learning (ML) holds promise to develop models to integrate complex, multimodel and high dimensional data from multiple data acquisitions to improve diagnosis, treatment and overall healthcare quality.

This project aims to develop appropriate flexible data harmonisation frameworks to link research and health data in any digital format allowing physicians and clinicians to study a disease system in a holistic manner, and test new data-driven hypotheses to develop early detection, diagnosis, and personalised treatment of disease.

Project aims

Define the aims of the project

1. Preparing catalogue of data related to disease of interest
2. Development of a robust bioinformatics pipeline to study genomic elements utilising 'multi-omics' information.
3. Integration of genomic data with health records and other possible information such as from IoT devices using analytics tools based on machine learning
4. AI and ML models for early diagnosis of diseases and identification of therapeutic targets

Expected outcomes

Highlight the expected outcomes of the project

- Novel biomarker discovery in human disease.
- Integrative Machine Learning based framework for modeling disease of interest.
- New predictive disease models suitable for personalized medicine.

How will the project address the Goals of the above Themes?

Describe how the project will address the goals of one or more of the 6 Themes listed above.

The project will combine the knowledge of Artificial Intelligence and Machine Learning with that of disease genetics to discover novel biomarkers and therapy targets.

Capabilities and Degrees Required

List the ideal set of capabilities that a student should have for this project. Feel free to be as specific or as general as you like. These capabilities will be input into the online application form and students who opt for this project will be required to show that they can demonstrate these capabilities.

- Candidates with a degree in Computer Science or Bioinformatics or related field will be preferred.
- Demonstrated advanced knowledge of computer programming and data exploration, analysis and visualization : preferred
- Competencies in computational biology and data analysis algorithms: preferred
- Prior exposure to machine learning frameworks (such as tensorflow, keras, PyTorch) is desirable
- Expertise in computer science and engineering with experience in distributed systems and networking; software engineering: preferred
- Knowledge of human genetics and disease: preferred
- Knowledge of genomics and current high throughput genomic technology: preferred