Indian Institute of Technology Kanpur

Proposal for a new course

1. **Course No.:**
2. **Course Title:** Materials Informatics
3. **Credits: 3-0-0-27; Duration of Course: Full Semester (Total of ~40 Lectures)**
4. **Proposing Department: Materials Science and Engineering**
   Other Departments which may be interested in the proposed course: CHE, ME, MSP
5. **Proposing Instructor: Krishanu Biswas**

   Other Instructors interested in taking this course: Dr. Shikhar Misra, Dr. Shikhar Krishn Jha, Dr. Somnath Bhowmick, Dr. Rajdip Mukherjee, Dr. Shivam Tripathi, Dr. Tanmoy Maiti
6. **Course Details**

Objectives: In this course, we will learn about the impact of data science on materials engineering with emphasis on Materials Informatics. Although computer models of materials are widely used, but there is an increasing trend to data driven approaches. This data might be derived from high throughput computations as well as experiments. This is course is intended to provide knowledge on materials design, discovery with applications of modern material informatics tools and large-scale multiscale modeling—with the ultimate goal to speed up the design process and implement cost effective rapid discovery, prototyping and elucidating the decision-making process.

7.

Sl. No	Broad title	Topics	Lectures
Introductory Concepts			
1	**Introduction to the Course**	Understanding the need of use of materials informatics in Materials Science and engineering. The importance data and informatics in material design. Historical Perspectives and what do we expect from MI in materials design.	**1**
2	**Statistics & Machine learning; Data Handling**	Elementary Concepts in Statistics and Regression – types, Bayes' theorem & Naïve-Bayes; Probability based learning and classification Other machine learning tools - decision tree, Clustering, ANN etc., Data handling, Gradient Descent, its variants and hands-on implementation – getting an intuitive feel of Materials Informatics	**4**
3	**Data mining**	Data warehouses, database management systems, data pre-processing, algorithms, clustering, and text mining	**5**
Materials Data Science and Featurization			

4	**Materials Science Databases**	Different databases- thermodynamics, materials property, crystallographic databases, Data repositories, APIs for loading datasets available online, Synthetic data.	**4**
5	**Working across data modalities**	Analysis of images, voxel data, dynamical data, and graphs, language and symbolic methods and hybrid approaches; features in-depth discussion of materiomic databases, synthetic datasets, and data collection in materials development	**5**
6	**Data optimization**	Basic optimization tools: Bayesian optimization. Hyperparameter optimization, Gradient Descent and Gradient Descent and its variants etc.	**4**
6	**Visualization and data analytic methods**	Statistical methods, graphic rendering, virtual reality; as well as interpretable machine learning, approaches at mapping and classification – structure maps, processing maps, Principles of Featurization	**4**
7	**High throughput computation and experiments**	High throughput characterization of composition and structure, high throughput property management.	**4**

Structure of Materials and Microstructure Informatics			
8	**Structure prediction**	Structure of materials at multiple length scale, Vector representations, crystal structure graphs, The microstructure function and digital representation of the microstructure. Convolutional Neural Networks; Synthetic microstructures, Featurizing the processing parameters; correlating processing conditions with structure.	**4**
9	**Computational materials discovery/design**	Techniques of combinatorial synthesis, materials testing and measurements; Rapid characterization of materials data; extracting maximal information from characterization tools with examples and codes	**4**
10	**Some case studies and project on use of Materials Informatics on problems of materials engineering**	Relevant examples and case studies from a variety of fields and at distinct scales, from molecular to macro, including structural materials, additive manufacturing as well as nanotechnology	**2**
Total			**41**

Prerequisites: Basic Materials courses on structure, processing, characterization of materials; mechanical and functional Properties of materials

8. **Short summary of the course content:**

Material informatics is transforming the way materials are discovered, designed, developed, selected, and deployed. The course will begin with an introduction to statistical tools, machine learning, data handling and data mining. This is followed data science in materials; databases data repository. including modern document-oriented approaches, and then move onto a description of the fundamentals of data analytics, visualization and programming. We will discuss how material structure might be systematically described, taking an inspiration from crystallography, thermodynamic databases and bio-/cheminformatics. Materials models across the length scales will be introduced, with an emphasis on density functional based methods and molecular dynamics. The core computational task of optimization, machine learning, data mining will be discussed, along with its application to regression, machine learning and material structure prediction. We will further discuss whether materials science has entered the era of "big data", and explore data mining as a way to make sense of large amounts of data with high throughput experiments.

9. **Text book and Reference Books**

1. Tongyi Zhang, *An Introduction to Materials Informatics*, Springer, 2024 (Text book)
2. Keisuke Takahashi and Lauren Takahashi, *Materials Informatics and Catalysis Informatics: An Introduction*, Springer, 2024 (Text book)
3. Olexandr Isayev, Alexander Tropsha, Stefano Curtarolo (ed.); *Materials Informatics: Methods, Tools and Applications*, Wiley-VCH Verlag GmbH & Co., 2019 (Reference book)

Date: 08/10/20224 Proposer (Krishanu Biswas)

Date: DUGC/DPGC Convener: