

Course Code	Course Name	Credits
26CH612	SUPERVISED LEARNING IN CHEMISTRY	04

### Course Objectives

- To apply regression and classification techniques to chemical datasets.
- To analyze structure–property and structure–activity relationships using machine learning.
- To evaluate model performance, interpretability, and generalization in chemical systems.
- To integrate cheminformatics, computational chemistry, and data-driven modeling.
- To enable the design of predictive models for molecular properties, reactions, and materials.

### Learning Outcomes

Upon successful completion of this course, it is intended that a student will be able to:

- Understand the scope of supervised learning in chemical sciences.
- Identify appropriate data representations for chemical problems.
- Generate and select relevant chemical features.
- Reduce dimensionality while preserving chemical information.

### Unit 1 - Introduction to Machine Learning in Chemistry (12 Hrs.)

Overview of artificial intelligence in chemistry; role of supervised learning in chemical research; types of chemical data including spectroscopic data, molecular descriptors, and reaction datasets.

### Unit 2 – Feature Engineering (12 Hrs.)

Molecular descriptors; physicochemical descriptors (logP, molecular weight, TPSA); topological indices (Wiener and Randic indices); fingerprints including MACCS keys and Morgan/ECFP; quantum descriptors such as HOMO–LUMO gap and DFT-derived partial charges.

### Unit 3 - Domain-Specific Applications (10 Hrs.)

Drug discovery; QSAR/QSPR modelling; virtual screening pipelines; ADMET prediction; materials chemistry; property prediction including band gap and conductivity; catalyst design; reaction chemistry; yield prediction; reaction condition optimization; retrosynthesis prediction.

### Unit 4 – Tools for data analysis (14 Hrs.)

ChiSurf: Open-source Python-based software for analyzing TCSPC data, fluorescence correlation spectroscopy (FCS), and FRET, capable of global fitting.

<sup>13</sup>C NMR-chemical shift and structure correlation, <sup>13</sup>C coupling constants, Solid state NMR, Magic angle spinning.

### Unit 5 - Origin pro/Origin tool in spectroscopy (12 Hrs.)

NMR shift prediction; IR and Raman spectral assignment; UV band calculation; full width at half maximum (FWHM); fluorescence studies; data plotting; peak fitting; complex analysis.

### Reference Books:

1. Todeschini, R., & Consonni, V. Molecular Descriptors for Chemoinformatics. Wiley-VCH, 2009.
2. Cherkasov, A. et al. QSAR Modeling: Where Have You Been? Where Are You Going To? Journal of Medicinal Chemistry, 2014.
3. Leach, A. R., & Gillet, V. J. An Introduction to Chemoinformatics. Springer, 2007.
4. Raschka, S., Liu, Y., & Mirjalili, V. Machine Learning with PyTorch and Scikit-Learn. Packt, 2022.
5. Keith, J. A. et al. Combining Machine Learning and Computational Chemistry for Predictive Insights into Chemical Systems. Chemical Reviews, 2021.

### Websites and eLearning Sources:

1. <https://www.originlab.com/index.aspx?go=Solutions/Applications/Spectroscopy>
2. <https://www.coursera.org/in/articles/data-analysis-tools>
3. [https://www.youtube.com/watch?v=O8OSlrgXBd8&list=PLeWSImvDbpleVJEkXIwSkWpRiA0\\_sX](https://www.youtube.com/watch?v=O8OSlrgXBd8&list=PLeWSImvDbpleVJEkXIwSkWpRiA0_sX)

**COs and Bloom's Taxonomy Mapping – 26CH612**

<b>Course Outcomes</b>	<b>On completing P.G. program the students will be able to</b>	<b>BTL</b>
<b>CO1</b>	Define and explain the fundamental concepts of machine learning and its role in chemical research, including different types of chemical data such as spectroscopic data, molecular descriptors, and reaction datasets.	K1, K2
<b>CO2</b>	Apply supervised learning methods to domain-specific chemical problems such as QSAR/QSPR modeling, virtual screening, ADMET prediction, material property prediction, and reaction optimization.	K3
<b>CO3</b>	Analyze and interpret experimental and computational chemical data using specialized tools such as ChiSurf for TCSPC, FCS, and FRET studies.	K4
<b>CO4</b>	Evaluate spectroscopic and fluorescence data using OriginPro/Origin tools for peak fitting, spectral assignment, UV band calculations, and complex data analysis.	K5
<b>CO5</b>	Design and develop machine learning workflows for solving real-world problems in chemistry, including predictive modeling and data-driven decision-making.	K6

**BTL K1 and K2 – remembering and understanding, K3- Applying, K4 – Analyse, K5- Evaluate and K6- Create**

**Relationship Matrix – 26CH612**

<b>Course Outcomes</b>	<b>Programme Outcomes (POs)</b>						<b>Programme Specific Outcomes (PSOs)</b>					<b>Mean Score of Cos</b>
	<b>PO1</b>	<b>PO2</b>	<b>PO3</b>	<b>PO4</b>	<b>PO5</b>	<b>PO6</b>	<b>PSO1</b>	<b>PSO2</b>	<b>PSO3</b>	<b>PSO4</b>	<b>PSO5</b>	
<b>CO1</b>	3	3	2	2	1.5	2	3	2	3	1	2	2.2
<b>CO2</b>	3	3	3	2	1	3	3	2	3	2	1.5	2.4
<b>CO3</b>	3	3	3	3	1	3	3	2	3	2	1.5	2.5
<b>CO4</b>	3	3	3	2	2	2	3	2	3	2	2	2.4
<b>CO5</b>	2	3	3	1	2	3	3	1	3	3	3	2.4
<b>Total</b>												2.4

Mean Score: 3- High, 2- Medium/Moderate, 1-Low

