TCEA

COMPUTATIONAL DRUG RECOMENDATION

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
B.TECH 7TH SEMISTER

CONTENT

INTRODUCTION

QSAR

Being able to predict biological activities of chemical structures by analysis of quantitative characteristics of structure features (quantitative structural activity/property relationships (QSAR)). The ‘physiological/biological activity’ Φ expressed as a function of the chemical structure C.

$$\Phi = f(C)$$

PROBLEM STATEMENT

- Analyze large amounts of data from different sources, such as medical records, drug reviews, clinical trials, and more.
- Suggest the most suitable drugs for a patient's condition, based on their needs, preferences, and feedback.
- Reduce the time and cost of drug development and delivery, by finding new insights and optimizing the process.
- Improve the quality and safety of drug recommendation, by detecting potential side effects, interactions, or adverse reactions.
- Enhance the accessibility and availability of drug recommendation, especially in remote areas where doctors are not available.

DEVELOPMENT METHODOLOGY

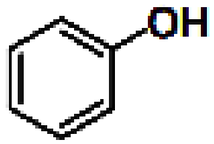
The construction of QSAR/QSPR model comprises of 3 main steps:

1. Description of molecular structure.
2. Multivariate analysis for correlating molecular descriptors with observed activities/properties.
3. Data preprocessing and statistical evaluation.

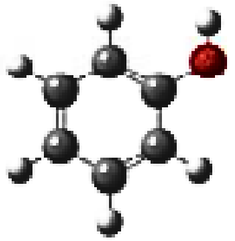
A molecular descriptor is a structural or physicochemical property of a molecule or part of a molecule.

Molecular Structures

OC1=CC=CC=C1 — 1D



2D



3D

Molecular Descriptors

Constitutional
Electronic
Geometrical
Hydrophobic
Lipophilicity
Solubility
Steric
Quantum Chemical
Topological

Data Pre-Processing

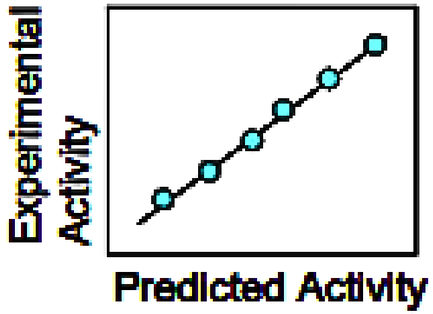
Normalization
Standardization
Feature Selection
Outlier Detection

Multivariate Analysis

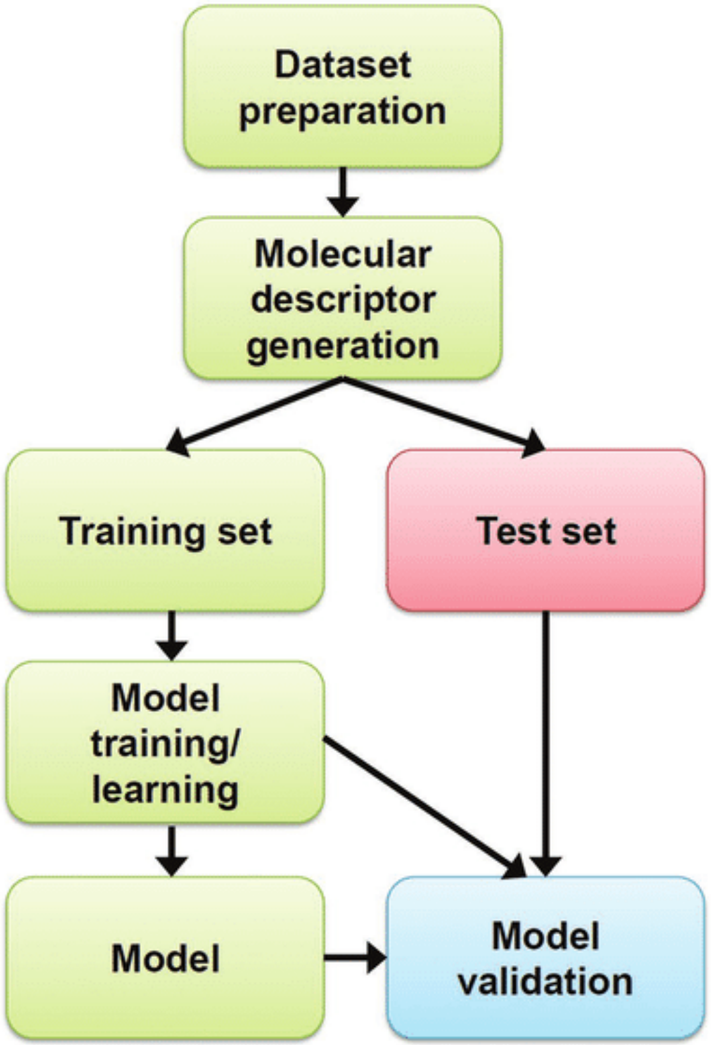
Multiple Linear Regression
Self-Organizing Map
Principal Component Analysis
Partial Least Squares
Neural Network
Support Vector Machine

Statistical Evaluation

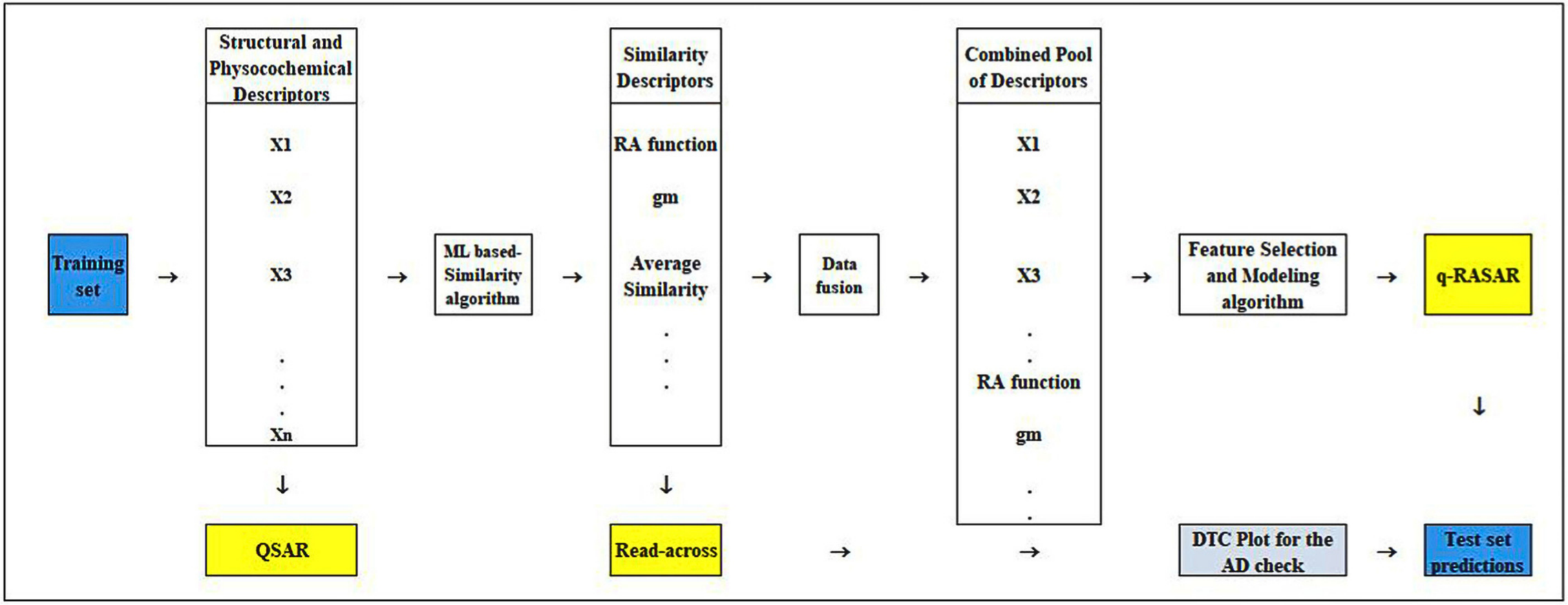
R
R^2
Q^2
MSE
RMSE



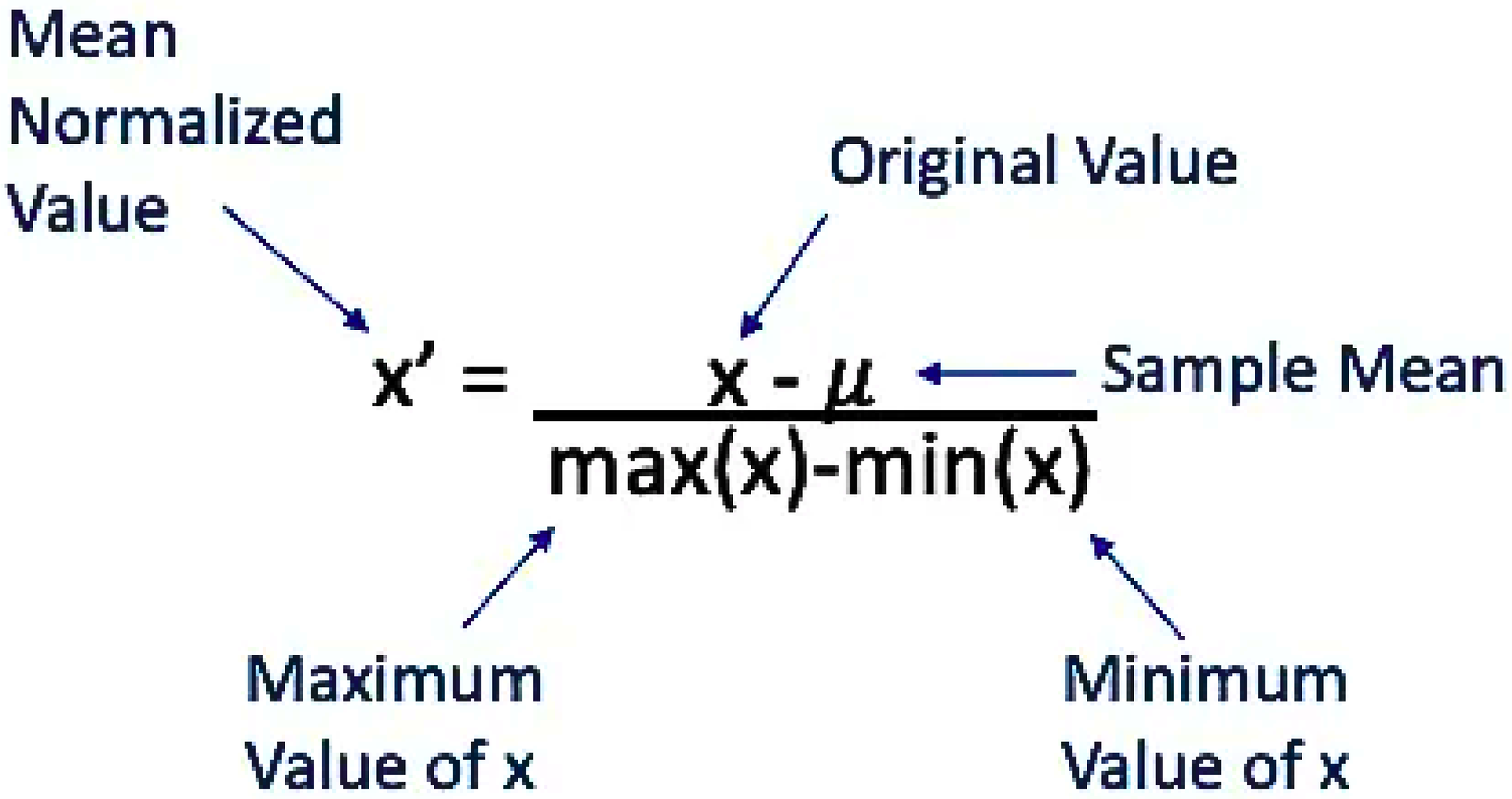
MODEL DESIGN



DATASET DESIGN



NORMALISATION



PRIMARY DATASET

The main sources of health statistics are surveys, administrative and medical records, health care claims data, vital records, surveillance, disease registries, grey literature and peer-reviewed literature.

SECONDARY DATASET

ChEMBL

13,377
Targets

1,950,765
Distinct compounds

15,996,368
Activities

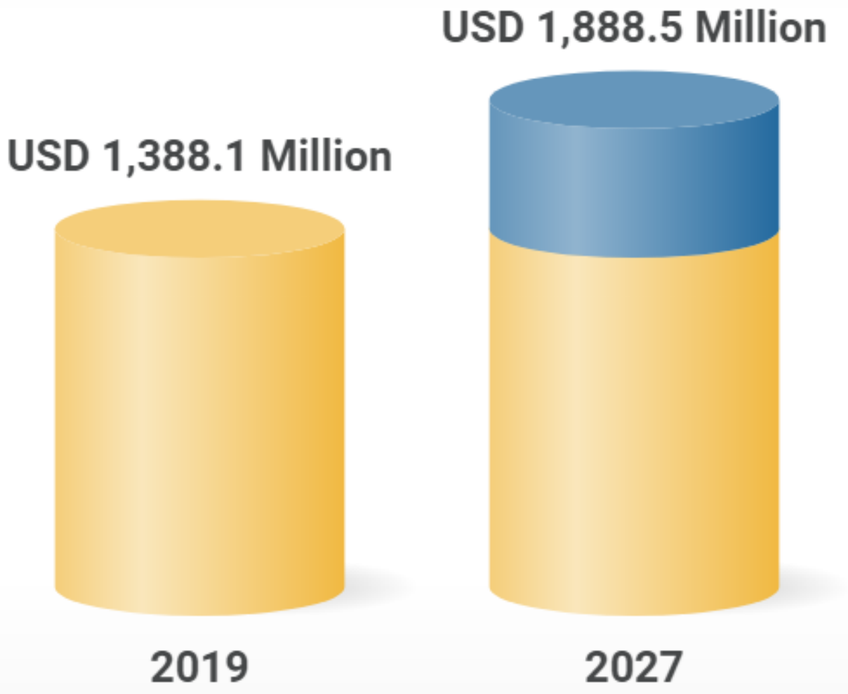
76,076
Publications

57
Deposited Datasets

SCOPE

Quantitative Structure-Activity Relationship (QSAR) Market

Market forecast to grow at a CAGR of 3.9%



https://www.researchandmarkets.com/reports/5308779

RESEARCH AND MARKETS
THE WORLD'S LARGEST MARKET RESEARCH STORE

THANK YOU