" 건강한 환경   행복한 미래 "

Various methods for Environmental data handling (Statistical Model, Machine Learning Algorithm)

국립환경과학원 영산강물환경연구소
연구사 황성윤

contents

01 Introduction

1. Environmental Data Handling (Water Environment Information System)

1. Environmental Data Handling (Water Environment Information System)

1. Introduction

1. Environmental Data Handling

Table 1b. Varimax rotated components

	1	2	3	4	5	6	7	8	9	10	11	12
Variance Explained by Rotated Components												
	2.470	2.044	1.043	1.014	1.022	1.015	1.033	0.968	0.742	0.369	0.145	0.135
Percent of Total Variance Explained												
	1	2	3	4	5	6	7	8	9	10	11	12
	20.585	17.037	8.690	8.452	8.517	8.457	8.610	8.070	6.181	3.073	1.206	1.121
Rotated Loadings												
	1	2	3	4	5	6	7	8	9	10	11	12
Q	-0.293	0.868	-0.093	0.008	0.046	0.174	0.060	0.002	0.168	-0.029	-0.003	0.295
T	0.942	-0.070	0.010	0.088	0.090	-0.047	0.015	-0.173	0.008	0.068	0.235	0.021
pH	0.059	-0.021	-0.020	0.038	0.994	0.006	-0.065	-0.000	0.044	0.012	0.004	0.004
EC	0.725	-0.316	0.053	0.064	0.046	-0.024	0.032	-0.104	-0.072	0.589	0.018	0.001
SS	0.019	0.924	-0.018	0.100	-0.078	0.209	0.032	0.050	0.163	-0.085	-0.000	0.217
MA1	0.032	-0.285	0.054	0.013	-0.007	-0.950	-0.017	-0.011	-0.111	0.008	0.002	0.004
Cl	0.129	-0.062	0.958	-0.046	-0.021	-0.054	0.215	0.044	0.089	0.014	-0.008	0.005
NH_3-N	0.068	0.076	0.229	0.133	-0.080	0.017	0.938	0.086	0.159	0.009	-0.004	0.002
NO_3-N	-0.140	-0.079	0.043	-0.971	-0.041	0.011	-0.119	-0.114	-0.009	-0.017	0.002	0.003
DO	-0.906	0.028	-0.175	-0.097	0.022	-0.009	-0.075	0.192	-0.006	0.077	0.299	.008
Pv	-0.016	0.482	0.151	0.014	0.083	0.182	0.262	0.020	0.796	-0.033	-0.000	0.006
BOD_5	-0.323	0.039	0.049	0.130	-0.000	0.012	0.088	0.930	0.015	-0.031	0.004	0.001

Table 2a. Principal components with an eigenvalue less than1

	1	2	3	4	
Latent Roots (Eigenvalues or Variances) Explained by Principal Components					
	3.481	2.456	1.549	1.162	
Percent of Total Variance Explained					
	1	2	3	4	
	29.007	20.467	12.905	9.683	
Component Loadings					**Communalities**
	PC1	PC2	PC3	PC4	
Q	0.813	0.337	-0.262	-0.017	0.843
T	-0.734	0.539	-0.238	0.075	0.891
pH	-0.088	0.061	-0.246	0.472	**0.294**
EC	-0.821	0.312	-0.067	0.079	0.782
SS	0.644	0.575	-0.264	-0.003	0.815
Mal	-0.463	-0.380	0.357	0.083	**0.493**
Cl	-0.183	0.372	0.646	-0.349	0.710
NH_3-N	0.097	0.565	0.641	-0.037	0.740
NO_3-N	0.029	-0.378	-0.163	-0.775	0.770
DO	0.659	-0.623	0.113	0.067	0.839
Pv	0.524	0.675	0.061	-0.120	0.748
BOD_5	0.467	-0.176	0.539	0.421	0.716

1. Tool for data analysis

2. R

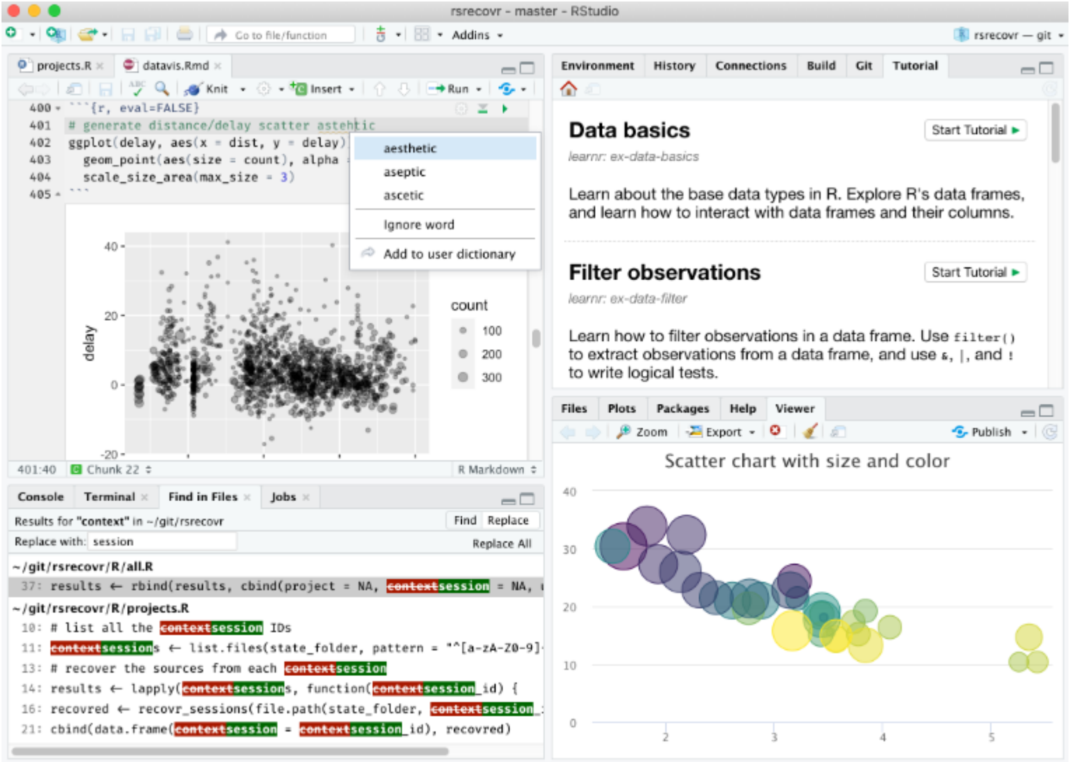
(1) 통계 계산과 그래픽을 위한 프로그래밍 언어이자 소프트웨어 환경

(2) 1960년대와 1970년대 Bell 연구소에서 개발된 S language에 기반을 두고, 1990년대 중반 뉴질랜드 오클랜드 대학의 로스 이하카와 로버트 젠틀맨에 의해 시작

(3) GPL 하에 배포되는 공개 소프트웨어로 누구나 자유롭게 이용할 수 있으며 자발적 기여자들에 의해 지속적으로 개발 중임

(4) 대표적 특징
1) 효율적이고 편리한 데이터 조작 및 처리
2) 데이터를 다양한 그래프로 표현해주는 데이터 시각화 기능
3) 통계 분석 및 데이터 마이닝 알고리즘 수행
4) 간단하며 효과적인 프로그래밍 언어로서의 기능

R Programming

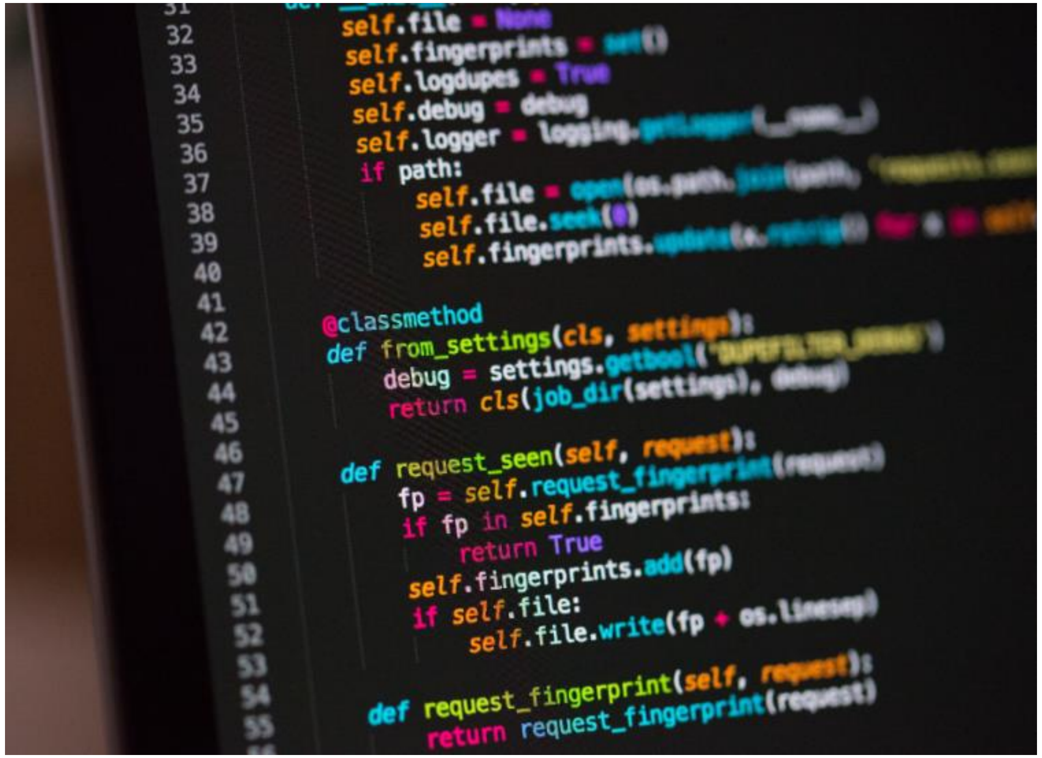


3. Python

(1) 1991년 네덜란드계 소프트웨어 엔지니어인 귀도 반 로섬에 의해 개발된 고급 프로그래밍 언어

(2) 플랫폼에 독립적이며 객체 지향적, 동적 타이핑 대화형 언어

(3) 강력한 라이브러리와 풍부한 생태계를 통해, 데이터를 수집하고 분석하며 시각화할 수 있음



(4) 대표적 특징
1) 문법이 쉬워 빠르게 배울 수 있음
2) 무료이지만 강력한 데이터 처리기능 보유
3) 간결한 형식의 코딩
4) 개발 속도가 빠르며 인공지능 기반 model에 특화되어 있음

1. Big data for environmental science

The EPA and Public Health

One of the biggest areas in the US for unifying big data with environmental science is public and environmental health **(16)**. Already, we've seen improvements in the monitoring and mitigation of toxicological issues of industrial chemicals released into the atmosphere. Monitoring has always used the tried and tested methods such as localized environmental sampling, but now we can process such data through computational methods, the result is more accurate, more up-to-date, faster produced, with more analytical information to allow experts to make an informed decision. Big Data allows for high throughput (more resources, a longer period of time), combined data sets (bringing together multiple, otherwise seemingly disparate data sets) and meta-analysis (studies that are the compilation of existing studies to create a more thorough and hopefully accurate picture), and deeper analysis of the results produced from these studies.

EPA is presently using such data acquired through Big Data Analytics to synthesize more accurate predictions for areas where data either does not exist or is difficult to acquire. Also, researchers can identify gaps in the data and potential vulnerabilities in the system and process of investigation. Overall, this mitigates the problems and enhances data for better decision making for public health concerns. They are now working with NCDS (National Consortium for Data Science) to identify current challenges that they hope to address through big data science **(16)**.

For Geographic Data

Few tools have proven as useful to so many environmental sciences as the map. From simple cartography for naval navigation, geographic surveying, to modern uses for Geographic Information Systems (databases of data sets from which we can produce digestible maps and create visually striking imagery for an intended audience), GIS thrives on Big Data. Much of GIS strength lies in its ability to consolidate, utilize and present statistical data. The more data you have from a geographic area, the better the quality of the output and the more informed the decision making is likely to be. Its biggest contribution (so far) seems to be in spatial analytics, and that's good news for GIS technicians and for those people charged with making decisions based on the outputs of their data.

One example is in disaster and emergency relief **(17)**. As recently as 2017, a researcher showed in a seminal study that it would be possible in future to parse textual references to GIS databases for up-to-the-minute problem areas currently suffering from tsunamis, flooding, and earthquakes. This would not have been possible before due to the sheer intensity of cross-referencing requirements. Satellite data and aerial imagery have already informed GIS in disaster management, with Hurricane Katrina being one of the first and best-known choices in using the technology. In future, Big Data will further enhance its efficacy.

Further, the EPA is using geographic data to inform research into public health through the Environmental Quality Index **(16)**. Big Data is informing a number of areas and bringing them together in the most comprehensive analysis of its kind examining air, water, and dry land, and the built environment and socio-economic data **(18)**. It is expected that this information will inform public health decisions and allow for medical research into health disparities of child mortality and poverty.

Reference : https://www.environmentalscience.org/data-science-big-data

1. Big data for environmental science

Climate Change and Planetary Monitoring

In 2013, the UK government announced large-scale investment in Big Data infrastructure for science, particularly in the environmental sector. Of particular note to global research was a commitment to maintaining funding for a program called CEMS (Climate and Environmental Monitoring from Space) **(19)**. This allowed for the creation of larger databases to cope with the upcoming Big Data revolution and to allow research partner organizations to work with more data and produce more results. With a specific focus on climate change and planetary monitoring, CEMS storage removed the need to download enormous data sets while reducing the cost of access **(20)**. It provides the tools as well as the data, allowing for greater efficiency, sharing in the academic community, and providing resources once beyond the reach of many institutes due to budgetary restrictions alone. Along with Cloud data, this is now the standard globally for some of the world's top research institutes.

At the same time, one of the UK's top universities announced plans to open a Big Data center for environmental science research and analysis. It intends to bridge the "data gap" between those who research global environmental problems and those charged with making decisions to remedy such issues **(21)**. That's also at the core of the relationship between the US-based Lighthill Risk Network - an insurance representative organization - and the UK's Institute for Environmental Analytics - a data research organization. Working in partnership to see how big data can be applied to a variety of issues in risk management and natural disasters, particularly in light of increased frequency of erratic and extreme weather, Lighthill is now committed to developing *global* databases and making the business case for sharing data **(22)**. Such cross-government and partnerships between industry and government are working as shown with the previously discussed EPA programs and the EU-wide Copernicus Climate Change Service which recently went live.

Finally, there are immense implications for the uses of Big Data for climate modeling. As early as 2010, NASA was utilizing Big Data capture and storage for creating climate models to make the most accurate climate projection models yet **(30)**. It is estimated the agency stores as much as 32 petabytes of information for modeling purposes. Models thrive on enormous data sets, complex data and accumulated metadata. As far as the sciences are concerned, climate modeling could be the single most important area of academia for Big Data applications. Learn more about the history of climate change.

Reference : https://www.environmentalscience.org/data-science-big-data

02 Statistical Models

2. Statistical Models

1. Evaluation of algal species distributions and prediction of cyanophyte cell counts using statistical techniques

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-023-30077-8

RESEARCH ARTICLE

Check for updates

Evaluation of algal species distributions and prediction of cyanophyte cell counts using statistical techniques

Seong-Yun Hwang[1] · Byung-Woong Choi[2] · Jong-Hwan Park[1] · Dong-Seok Shin[3] · Won-Seok Lee[1] · Hyeon-Su Chung[1] · Mi-Sun Son[1] · Don-Woo Ha[1] · Kyung-Lak Lee[4] · Kang-Young Jung[5]

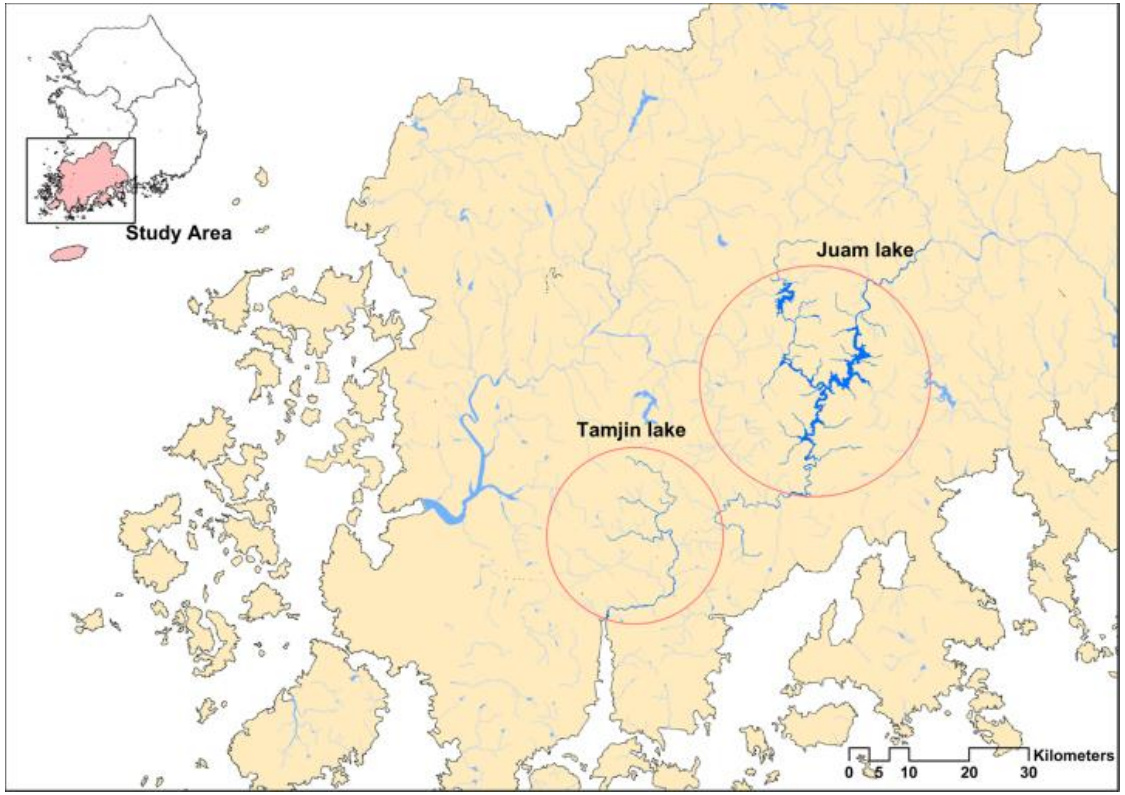
Received: 4 May 2023 / Accepted: 21 September 2023

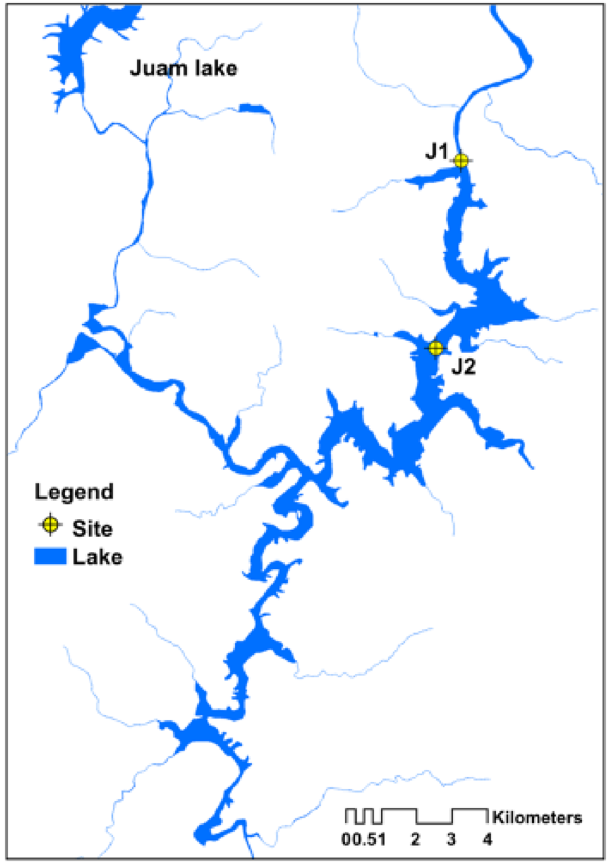
Abstract
Safe drinking water sources are crucial for human health. Consequently, water quality management, including continuous monitoring of water quality and algae at sources, is critical to ensure the availability of safe water for local residents. This study aimed to construct statistical prediction models considering probability distributions relevant to cyanophyte cell counts and compare their prediction performance. In this study, water quality parameters at Juam Lake and Tamjin Lake, representative water sources in the Yeongsan and Seomjin rivers, South Korea, were investigated. We used a water quality monitoring network, algae alert system, and hydraulic and hydrological data measured every 7 days from January 2017 to December 2022 from the Water Environment Information System of the National Institute of Environmental Research. Using data for 2017–2021 as a training set and data for 2022 as a test set, the performances of seven models were compared for predicting cyanophyte cell counts. Environmental factors associated with algae in water sources were observed based on the monitoring data, and a prediction model appropriate for the cyanophyte distribution was generated, which also included the risk of toxicity. The extreme gradient boosting with the random forest model had the best predictive performance for cyanophyte cell counts. The study results are expected to facilitate water quality management in various water systems, including water sources.

Keywords Water quality · Cyanophytes · Redundancy analysis · Statistical model · Random forest model · South Korea

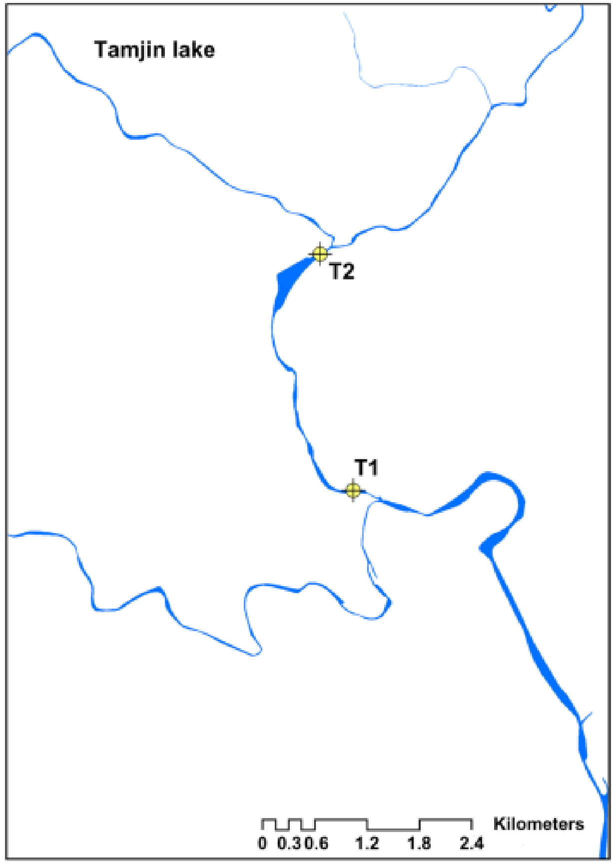
2. Statistical Models

1. Evaluation of algal species distributions and prediction of cyanophyte cell counts using statistical techniques





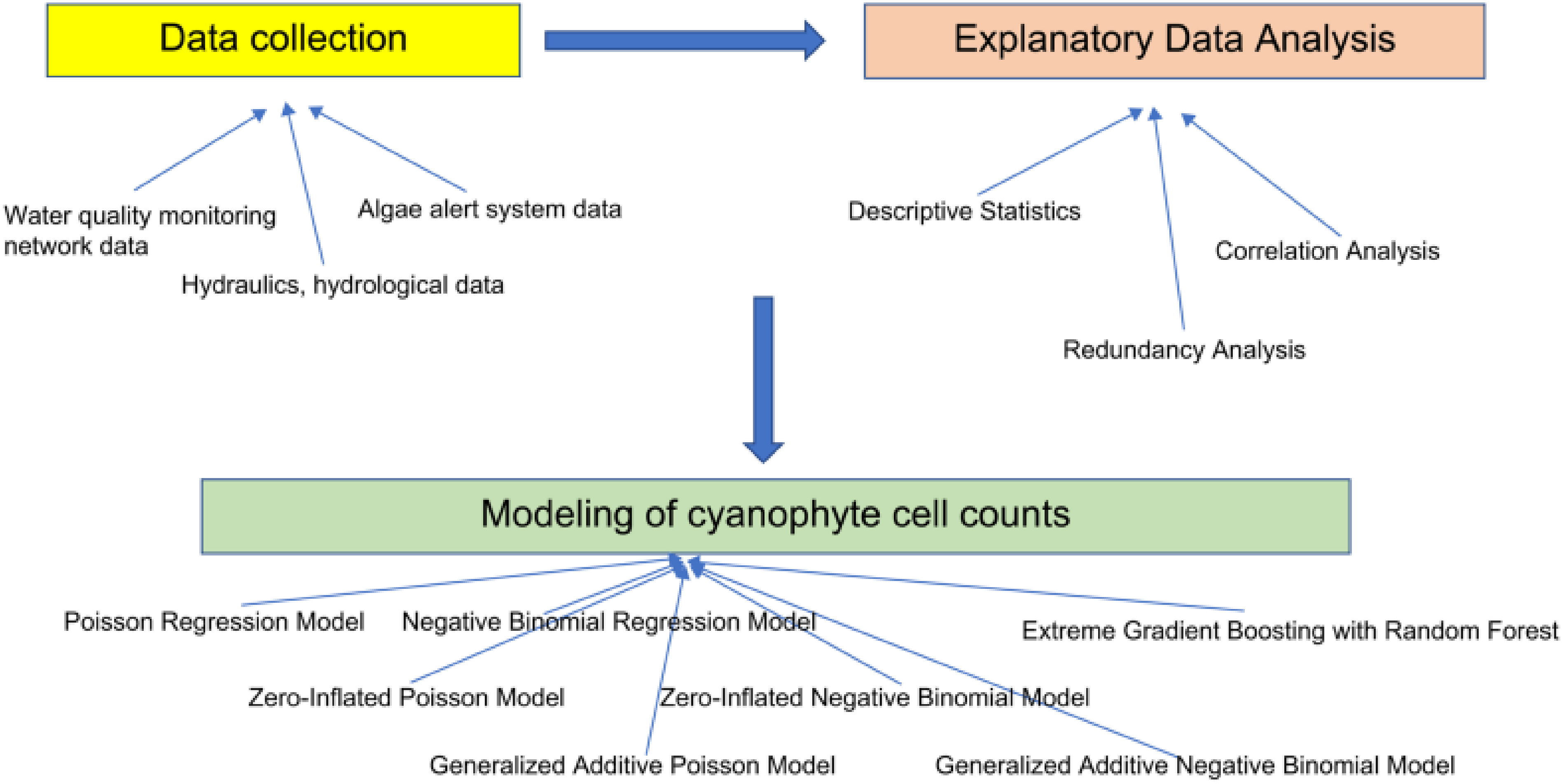
(a) Juam lake (J1, J2)



(b) Tamjin lake (T1, T2)

2. Statistical Models

1. Evaluation of algal species distributions and prediction of cyanophyte cell counts using statistical techniques



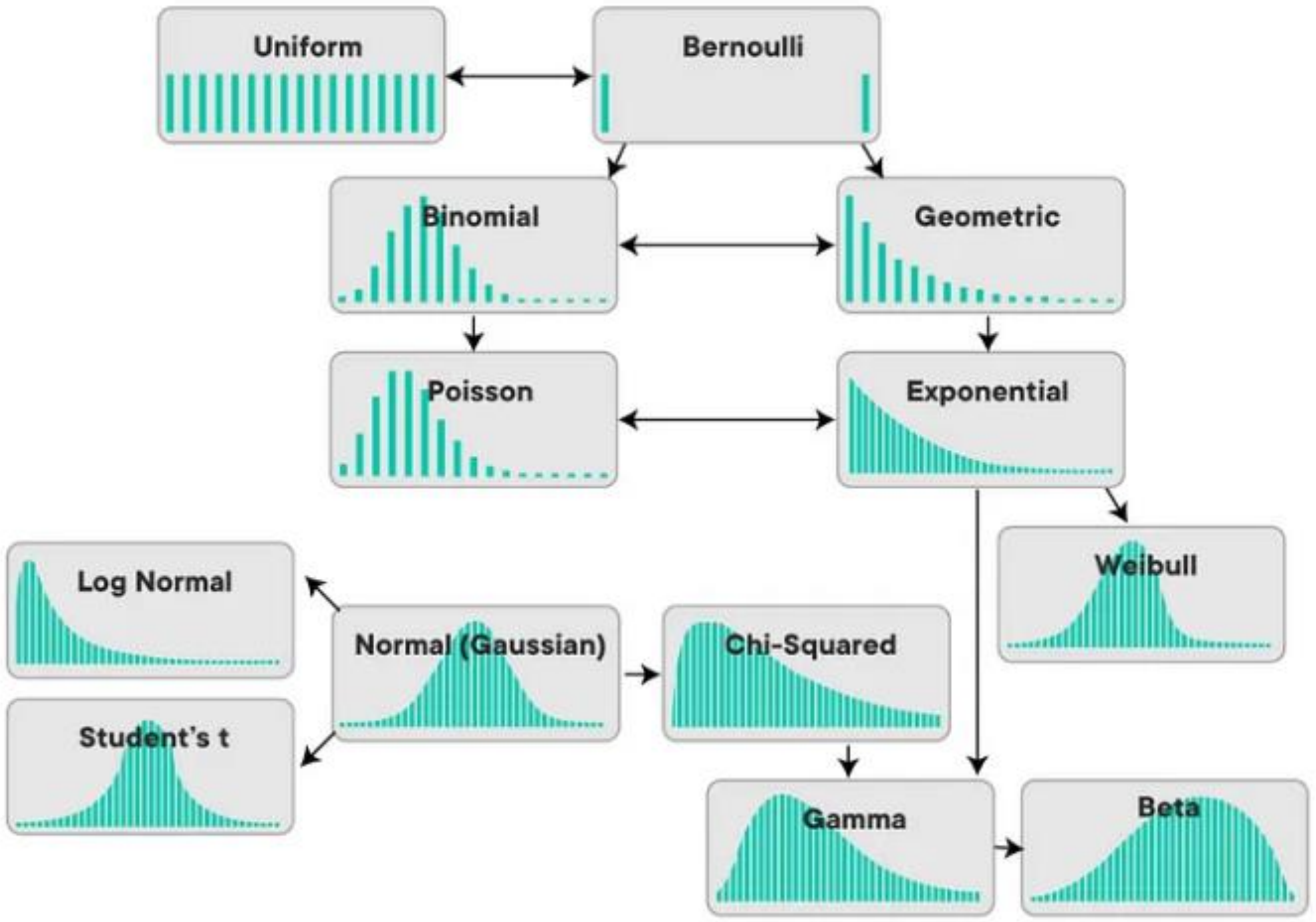
1. Evaluation of algal species distributions and prediction of cyanophyte cell counts using statistical techniques

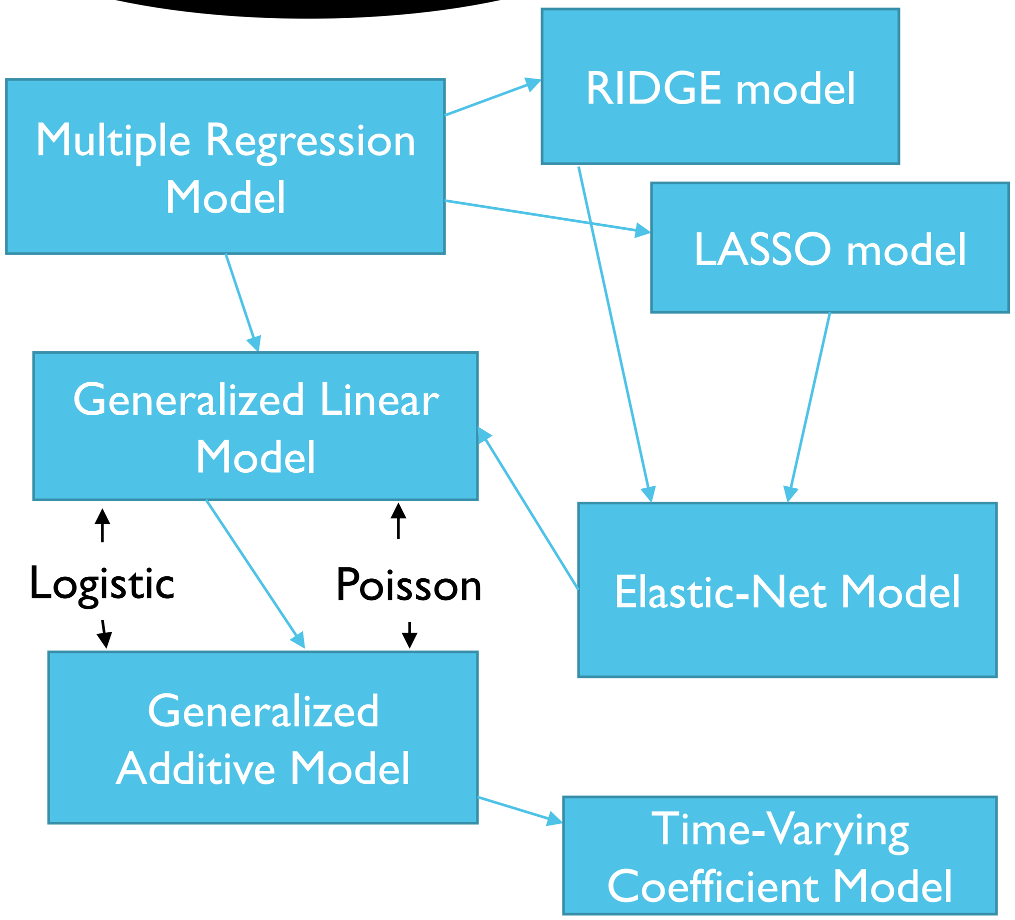
Table 1 Overview of variables

Response variable (count data)		Explanatory variables (continuous)	
		Water quality	Hydraulics, hydrological
Redundancy Analysis	Modeling of cell counts	BOD (mg/L)	Low water level (cm)
Cell counts of all algal species at the sampling site	Cyanophyte cell counts	COD (mg/L)	Inflow (cm^3/s)
		TN (mg/L)	Discharge (cm^3/s)
		TP (mg/L)	Reservoir (10,000 m^3)
		TOC (mg/L)	
		SS (mg/L)	
		EC (μS/cm)	
		pH	
		DO (mg/L)	
		Temperature (°C)	
		Turbidity (NTU)	
		Transparency (m)	
		Chl a (mg/m^3)	

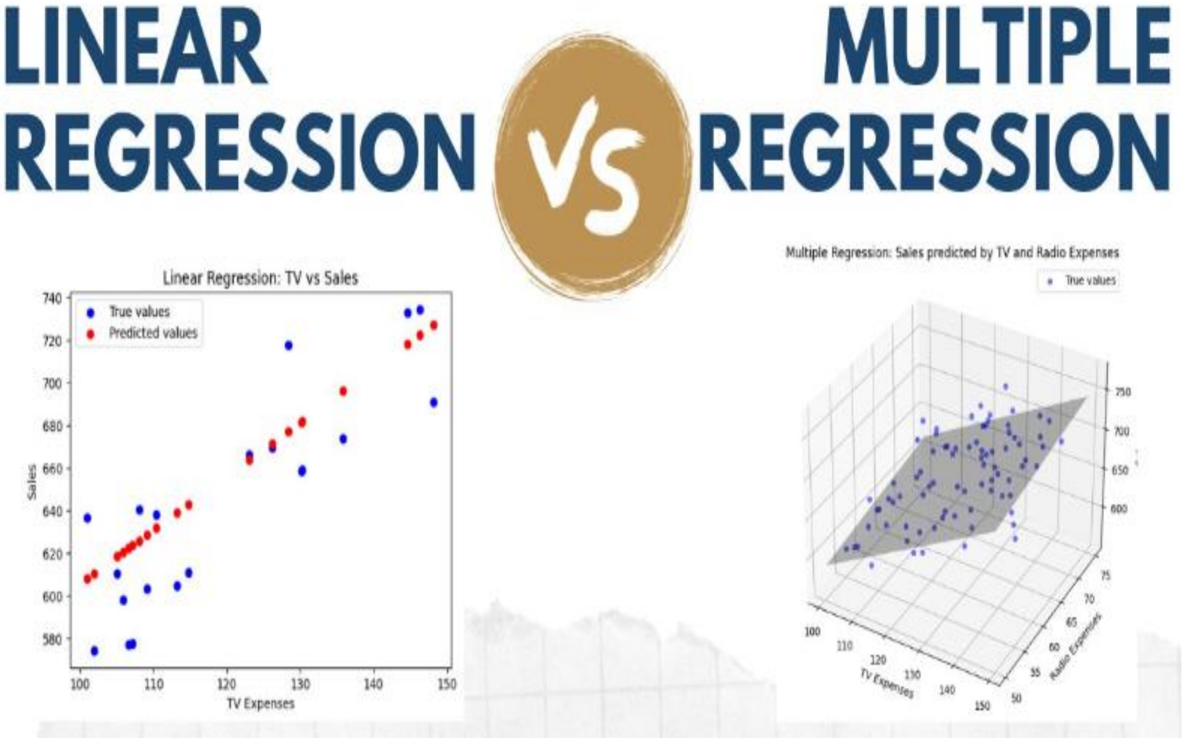
2. Various statistical models

Probability Distributions → Statistical Models





2. Various statistical models
1) Multiple Regression Model



One Predictor Model

$$Y = \underbrace{\beta_0 + \beta_1 x_1}_{\text{Nonrandom or Systematic Component}} + \underbrace{\varepsilon}_{\text{Random Component}}$$

Multiple Predictor Model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_q x_q + \varepsilon$$

Where

Y is the outcome value

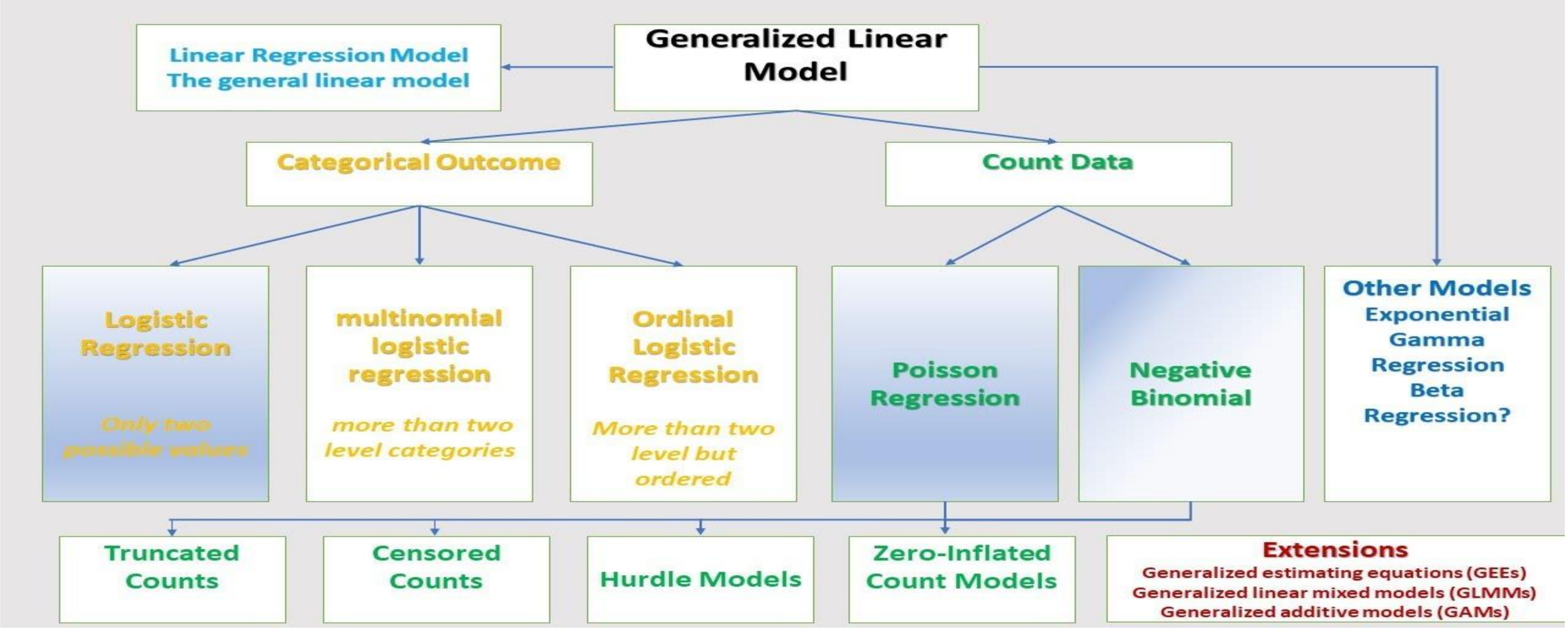
β_0 is the intercept

ε is the error aka residual

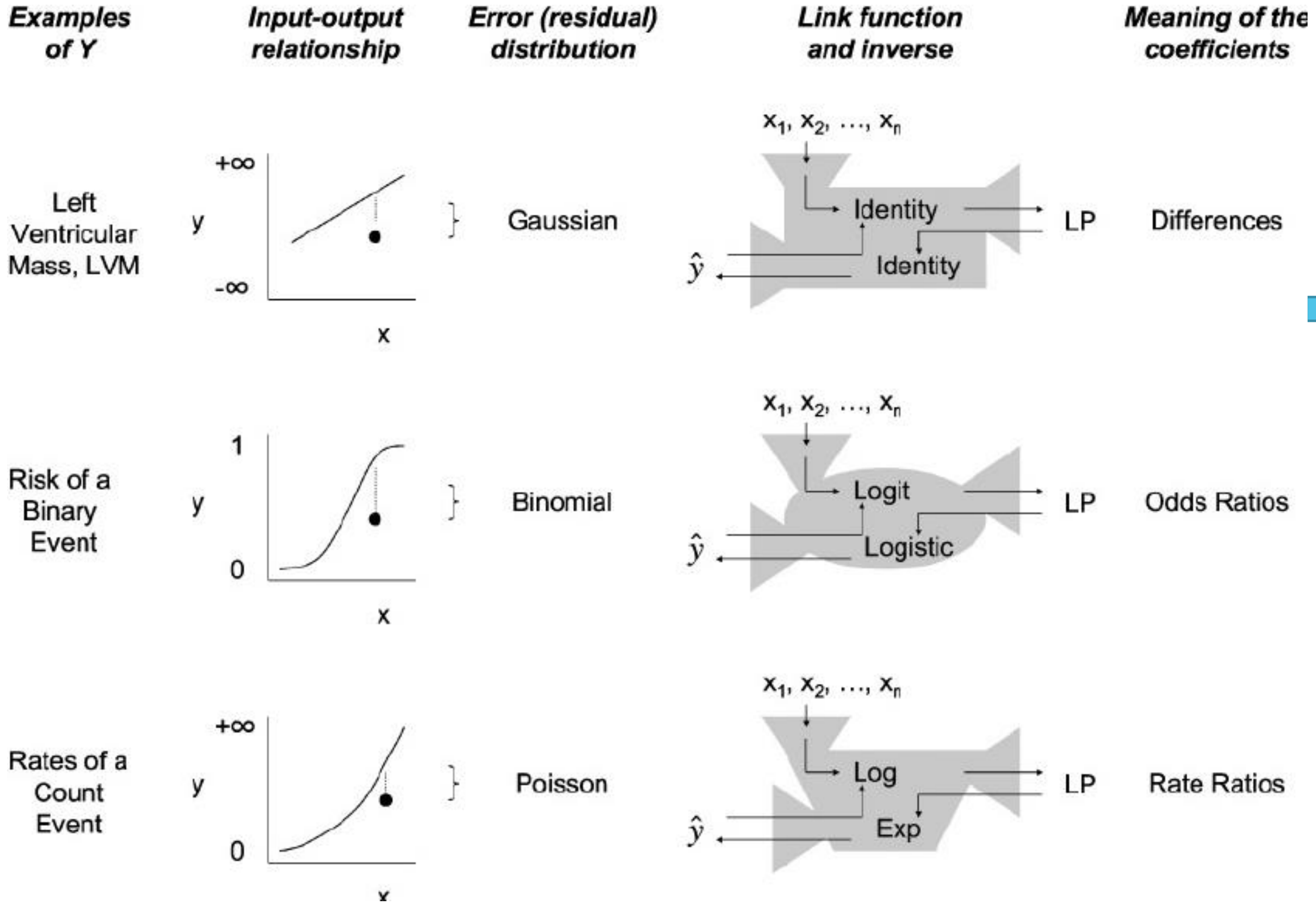
$x_{1..q}$ is the value of predictor variable

$\beta_{1..q}$ is the slope coefficient

2. Various statistical models
2) Generalized Linear Model



2. Various statistical models
2) Generalized Linear Model



Link function

Regression coefficients

Regression variables

$$g(\pi_i) = \sum_{j=0}^{p} \boldsymbol{\beta}_j \cdot \boldsymbol{x}_{ij}$$

where:

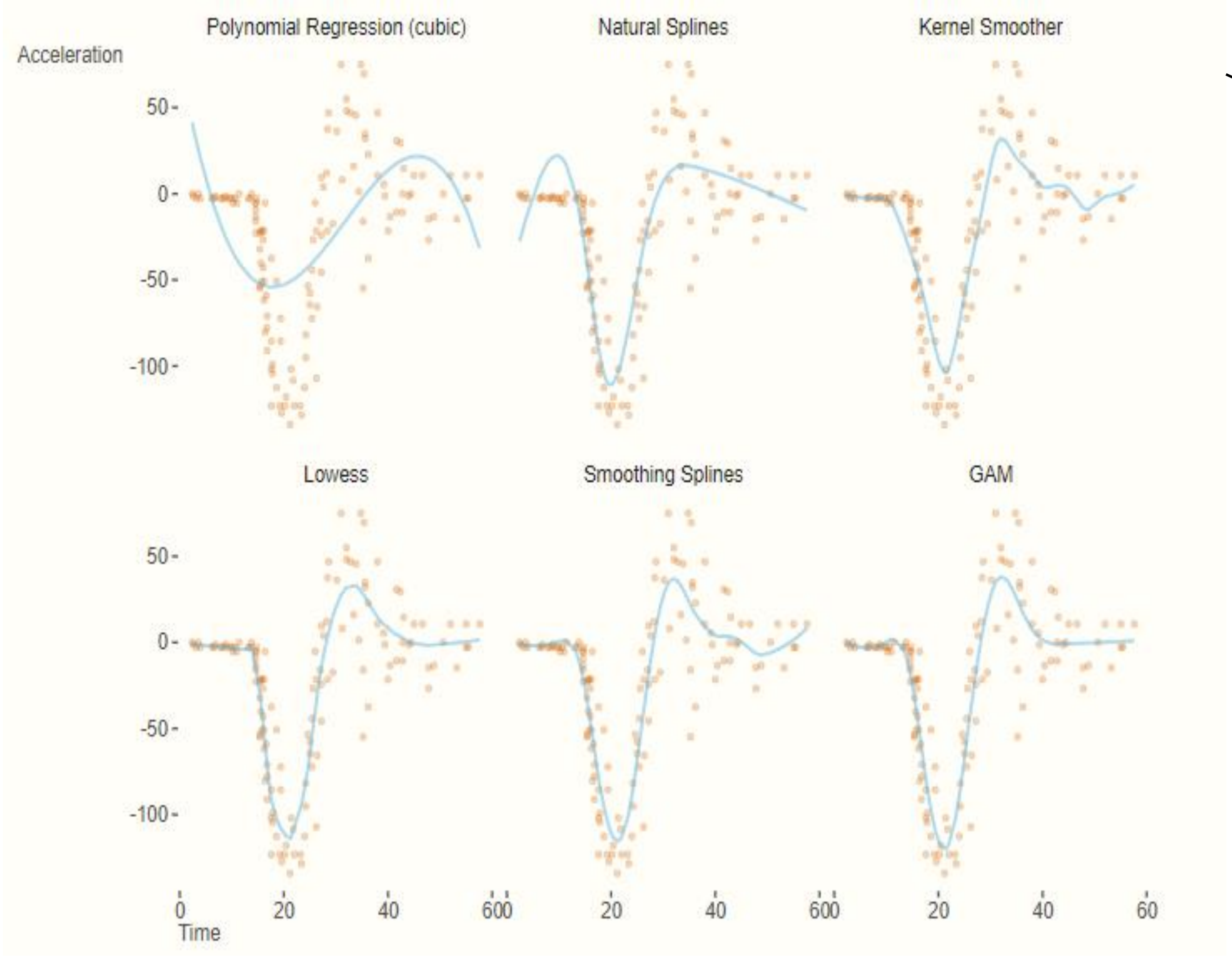
$$\pi_i = \underbrace{E(\boldsymbol{y} = y_i | \boldsymbol{X} = \boldsymbol{x}_i)}_{\text{Conditional expectation of } \boldsymbol{y} \text{ on } \boldsymbol{X}=\boldsymbol{x}_i}$$

Conditional mean a.k.a. conditional expectation

Conditional expectation of **y** on ***X*=*x*$_i$**

2. Statistical Models

2. Various statistical models
3) Generalized Additive Model



Model	Function
Linear Model	$y = b + w_1 x_1 + ... + w_D x_D$
Generalized Linear Model	$g(y) = b + w_1 x_1 + ... + w_D x_D$
Generalized Additive Model (GAM)	$g(y) = f_0 + f_1(x_1) + ... + f_D(x_D)$
GA^2M	$g(y) = \sum_i f_i(x_i) + \sum_{ij} f_{ij}(x_i, x_j)$
Full Complexity Model	$y = f(x_1, x_2, ..., x_D)$

03 Machine Learning Algorithms

3. Machine Learning Algorithms

1. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

water MDPI

Article

Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

Seong-Yun Hwang [1,†], Byung-Woong Choi [1], Jong-Hwan Park [1], Dong-Seok Shin [2], Hyeon-Su Chung [1], Mi-Sun Son [1], Chae-Hong Lim [1], Hyeon-Mi Chae [1], Don-Woo Ha [1] and Kang-Young Jung [3,*]

1 Yeongsan River Environment Research Center, National Institute of Environmental Research, 5, Cheomdangwagi-ro 208beon-gil, Buk-gu, Gwangju 61011, Republic of Korea; hsyliark@korea.kr (S.-Y.H.); bchoi628@korea.kr (B.-W.C.); thanks@korea.kr (J.-H.P.); jys7246@korea.kr (H.-S.C.); miza03@korea.kr (M.-S.S.); chaehong@korea.kr (C.-H.L.); chm2022@korea.kr (H.-M.C.); hahaha9909@korea.kr (D.-W.H.)
2 Freshwater Bioresources Culture Research Division, Nakdonggang National Institute of Biological Resources, 137, Donam 2-gil, Sangju-si 37242, Republic of Korea; sds8488@korea.kr
3 Education Planning Division, National Institute of Environmental Human Resources Development, 42, Hwangyeong-ro, Seo-gu, Incheon 22689, Republic of Korea
* Correspondence: happy3313@korea.kr; Tel.: +82-32-560-7795
† This author is the primary author of this study.

Abstract: South Korea's National Institute of Environmental Research (NIER) operates an algae alert system to monitor water quality at public water supply source sites. Accurate prediction of dominant harmful cyanobacterial genera, such as *Aphanizomenon*, *Anabaena*, *Oscillatoria*, and *Microcystis*, is crucial for managing water source contamination risks. This study utilized data collected between January 2017 and December 2022 from Juam Lake and Tamjin Lake, which are representative water supply source sites at the Yeongsan River and Seomjin River basins. We performed an exploratory data analysis on the monitored water quality parameters to understand overall fluctuations. Using data from 2017 to 2021 as training data and 2022 data as test data, we compared the dominant algal classification accuracy of 11 statistical machine learning algorithms. The results indicated that the optimal algorithm varied depending on the survey site and evaluation criteria, highlighting the unique environmental characteristics of each site. By predicting dominant algae in advance, stakeholders can better prepare for water source contamination accidents. Our findings demonstrate the applicability of machine learning algorithms as efficient tools for managing water quality in water supply source systems using monitoring data.

Keywords: water quality; Yeongsan River; Seomjin River; correlation analysis; self-organizing map; statistical machine learning algorithm; classification

check for updates

Citation: Hwang, S.-Y.; Choi, B.-W.; Park, J.-H.; Shin, D.-S.; Chung, H.-S.; Son, M.-S.; Lim, C.-H.; Chae, H.-M.; Ha, D.-W.; Jung, K.-Y. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea. *Water* **2023**, *15*, 1738. https://doi.org/10.3390/w15091738

Academic Editor: Guangyi Wang

3. Machine Learning Algorithms

1. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

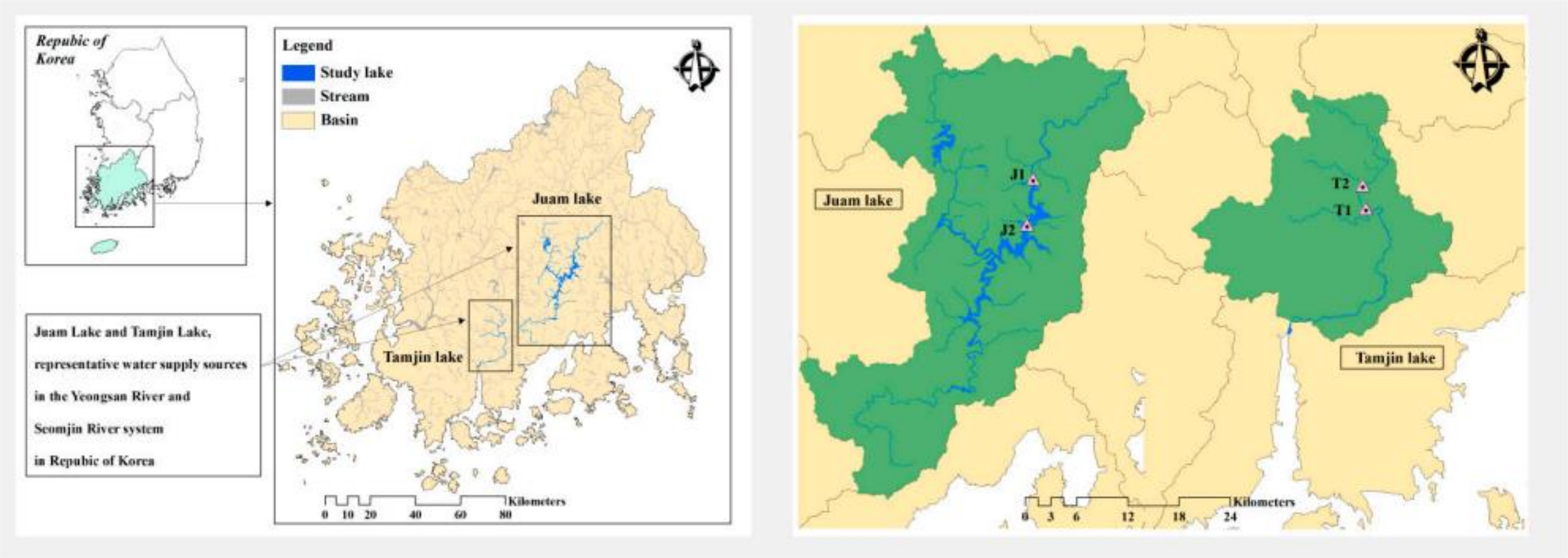


Figure 2. Sampling sites at Juam Lake and Tamjin Lake.

1. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

Table 1. Data variables used in this study.

Response Variable (Categorical)	**Explanatory Variables (Continuous)**	
Dominant Algae (Based on Total Cell Count)	**Water Quality**	**Hydraulic/Hydrological**
Cyanophytes Diatoms Chlorophytes Others	Biological Oxygen Demand (BOD), mg L^{-1} Chemical Oxygen Demand (COD), mg L^{-1} Total Nitrogen (TN), mg L^{-1} Total Phosphorus (TP), mg L^{-1} Total Organic Carbon (TOC), mg L^{-1} Suspended Solids (SS), mg L^{-1} Electrical Conductivity (EC), μS L^{-1} pH Dissolved Oxygen (DO), mg L^{-1} Temperature, °C Turbidity, NTU Transparency, m Chlorophyll a (Chla), mg m^{-3}	Low Water Level, cm Inflow Rate (Inflow), cms Discharge Rate (Discharge), cms Water Storage Capacity (Reservoir), 10,000 m^3

3. Machine Learning Algorithms

1. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

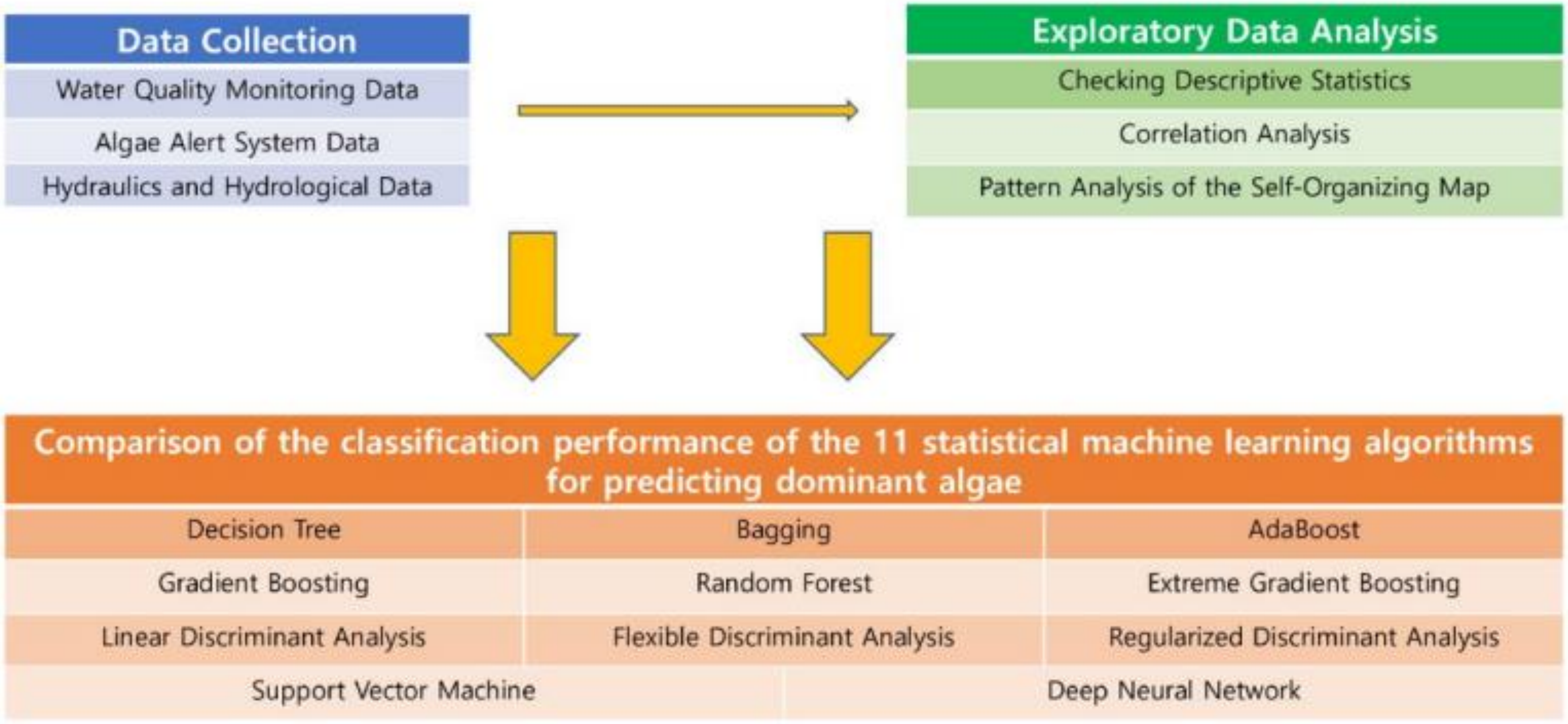


Figure 1. Methodological flowchart used in this study.

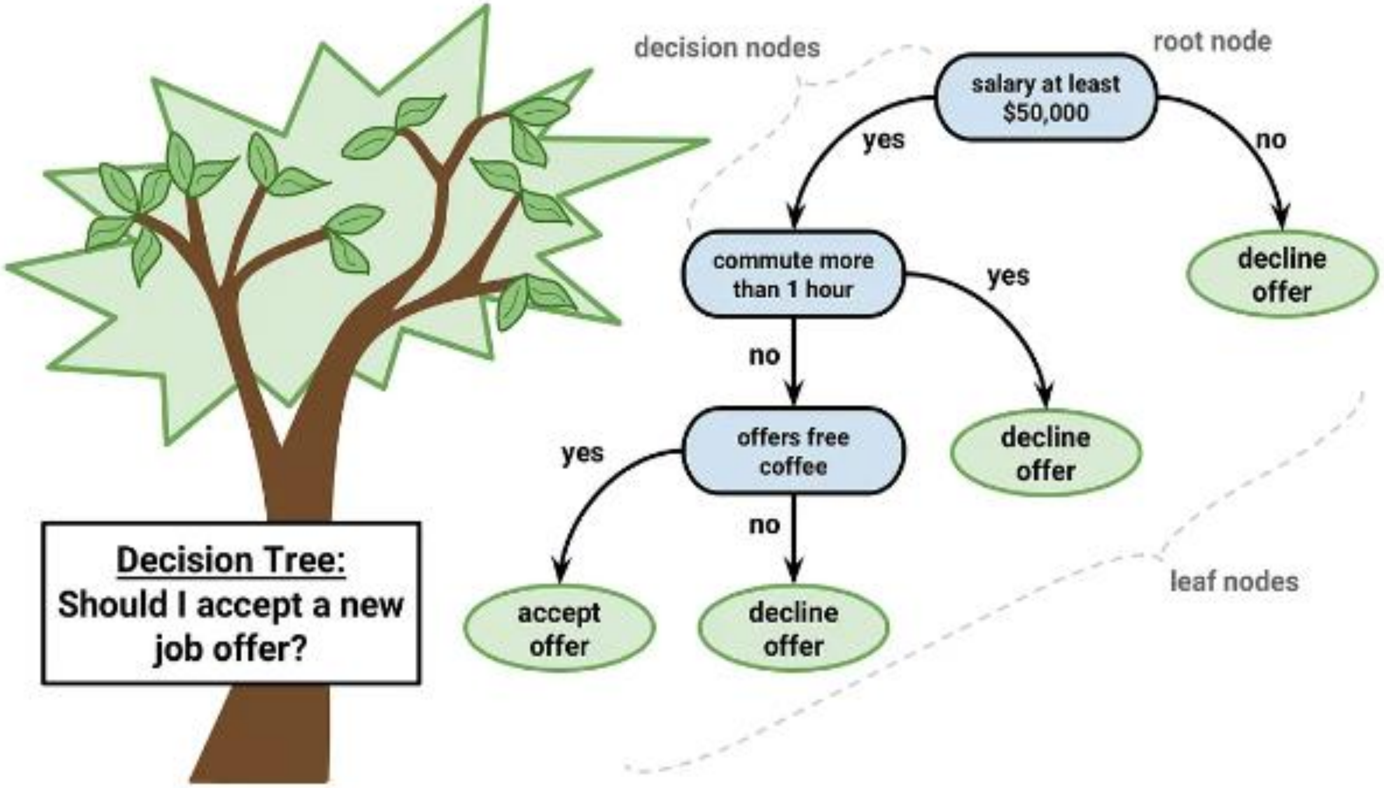
3. Machine Learning Algorithms

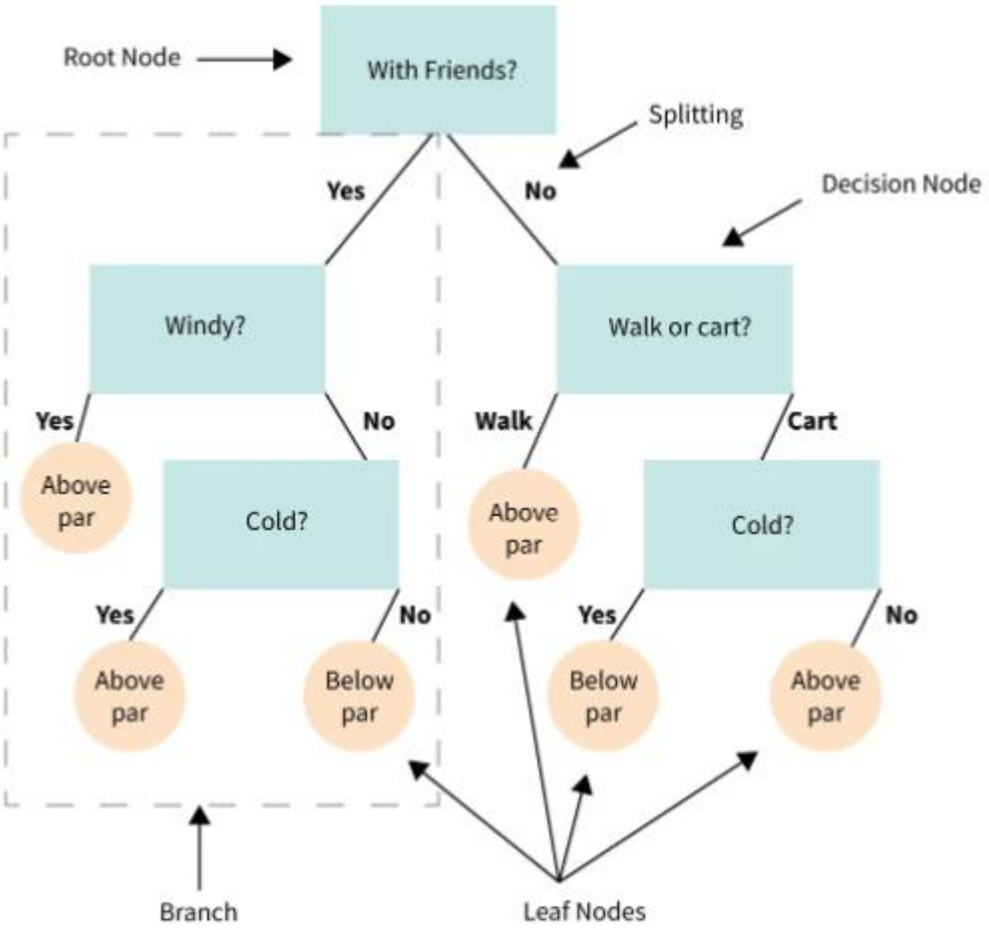
1. Evaluating Statistical Machine Learning Algorithms for Classifying Dominant Algae in Juam Lake and Tamjin Lake, Republic of Korea

Table 9. Result of dominant algal classification using 11 statistical machine learning algorithms (values in bold represent the criterion for which each algorithm shows the best performance, at each of the four sites).

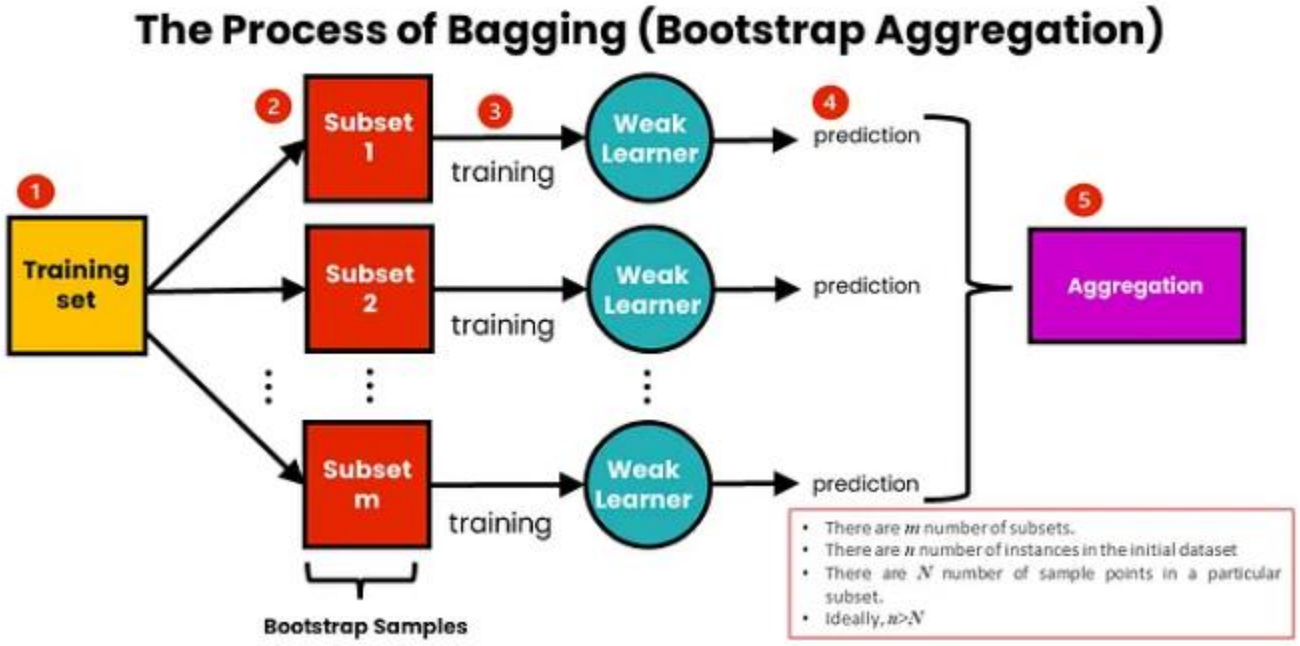
Site	Criterion	Algorithm										
		DT	Bag	Ada	GB	RF	XGB	LDA	FDA	RDA	SVM	DNN
J1	Accuracy	**0.7000**	0.6200	0.6000	0.5400	0.6200	0.6200	0.4000	0.4000	0.4200	0.6600	0.5800
	Weighted Sensitivity	**0.7000**	0.6200	0.6000	0.5400	0.6200	0.6200	0.4000	0.4000	0.4200	0.6600	0.5800
	Weighted Specificity	0.6239	0.6431	0.6949	0.7010	0.6699	0.6948	0.8791	0.8791	**0.9046**	0.6257	0.4200
	G mean	**0.6609**	0.6314	0.6462	0.6153	0.6445	0.6563	0.5930	0.5930	0.6164	0.6426	0.4936
J2	Accuracy	0.5800	0.5400	0.5400	0.5200	**0.6600**	0.5600	0.5800	0.5800	0.5400	0.6200	0.5400
	Weighted Sensitivity	0.5800	0.5400	0.5400	0.5200	**0.6600**	0.5600	0.5800	0.5800	0.5400	0.6200	0.5400
	Weighted Specificity	0.7620	0.7385	0.7046	0.7087	0.7179	**0.8067**	0.7131	0.7131	0.4600	0.6583	0.4600
	G mean	0.6648	0.6315	0.6168	0.6071	**0.6883**	0.6721	0.6431	0.6431	0.4984	0.6389	0.4984
T1	Accuracy	0.7551	0.8163	0.8367	0.8776	**0.9184**	0.7959	0.5918	0.5918	0.8367	0.8980	0.8367
	Weighted Sensitivity	0.7551	0.8164	0.8368	0.8775	**0.9184**	0.7960	0.5919	0.5919	0.8367	0.8980	0.8367
	Weighted Specificity	0.8641	0.7709	0.7762	0.7843	0.6834	0.8698	**0.8801**	**0.8801**	0.1633	0.7823	0.1633
	G mean	0.8078	0.7933	0.8059	0.8296	0.7922	0.8321	0.7218	0.7218	0.3696	**0.8382**	0.3696
T2	Accuracy	0.7551	0.7551	0.7551	**0.7755**	0.7551	0.7551	0.7143	0.7143	0.7551	0.7551	0.7551
	Weighted Sensitivity	0.7552	0.7552	0.7552	**0.7756**	0.7552	0.7552	0.7143	0.7143	0.7552	0.7552	0.7552
	Weighted Specificity	0.2448	0.2448	0.3043	0.3673	0.2448	**0.3698**	0.2439	0.2439	0.2448	0.2448	0.2448
	G mean	0.4300	0.4300	0.4794	**0.5337**	0.4300	0.5285	0.4174	0.4174	0.4300	0.4300	0.4300

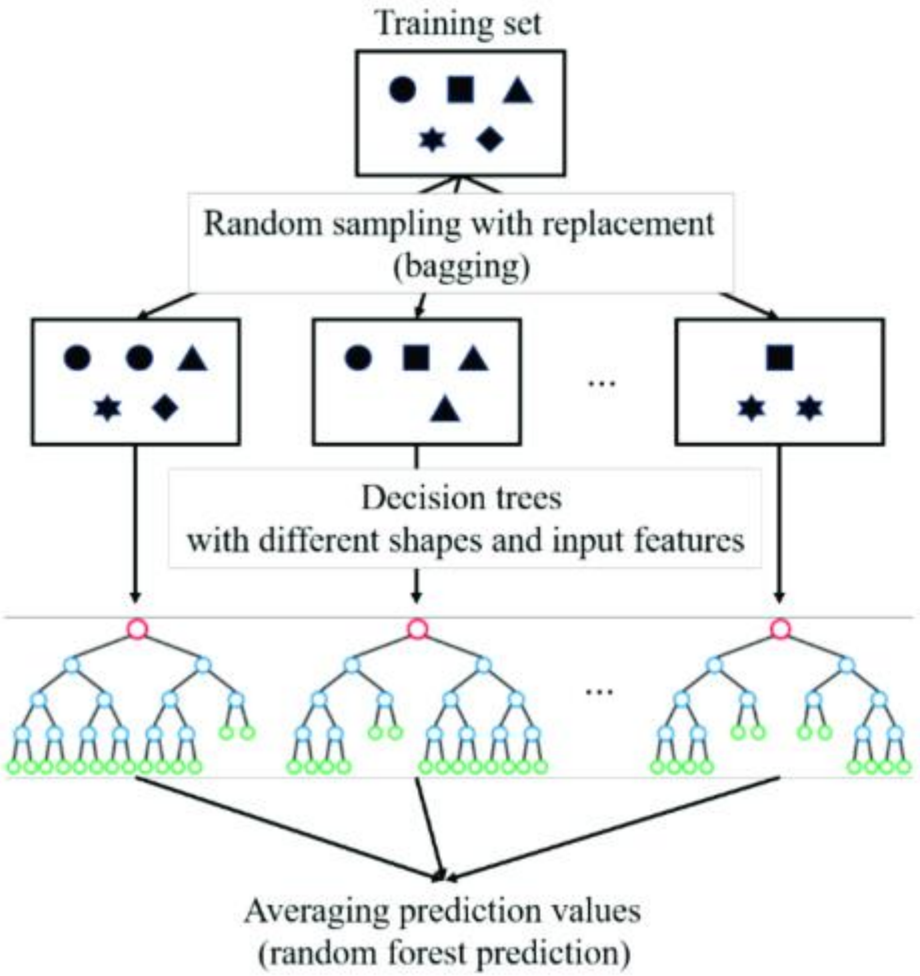
2. Various Machine Learning Algorithms
1) Decision Tree





2. Various Machine Learning Algorithms
2) Bagging and Random Forest





3. Machine Learning Algorithms

2. Various Machine Learning Algorithms
3) AdaBoost and Gradient Boosting

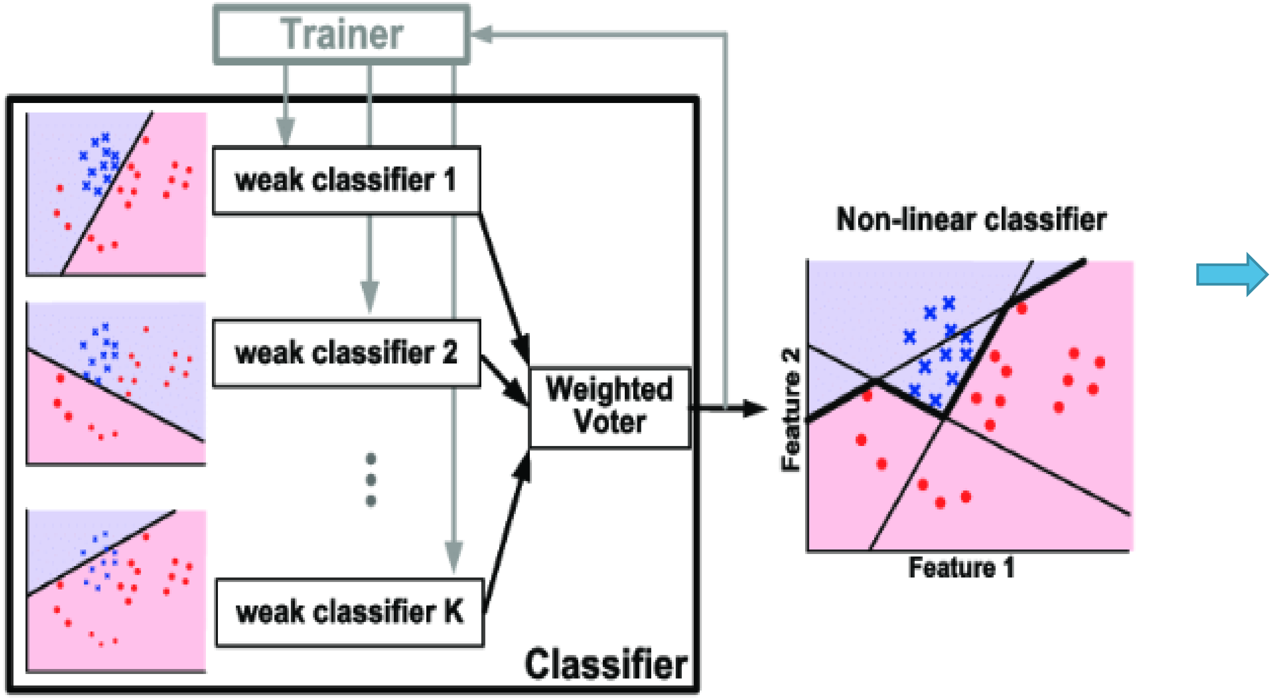
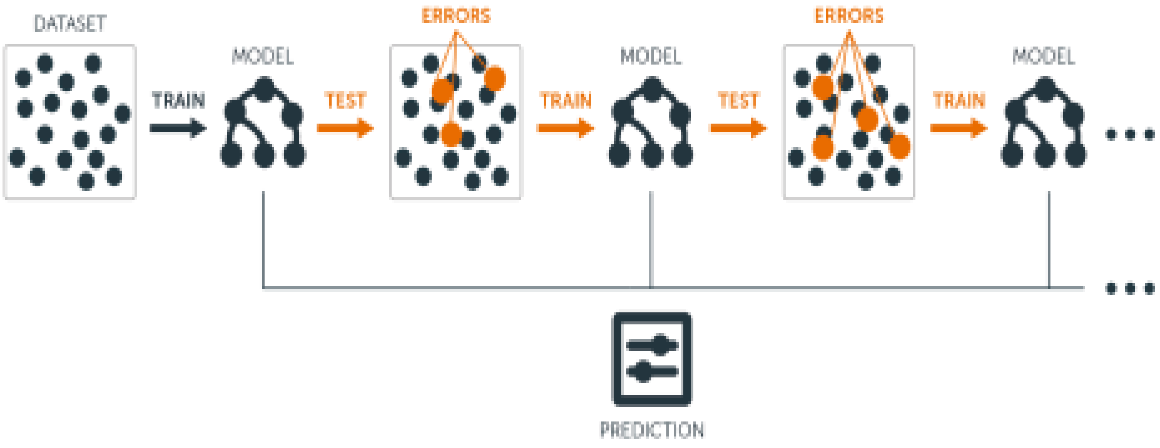


Illustration of AdaBoost algorithm for creating a strong classifier based on multiple weak linear classifiers.

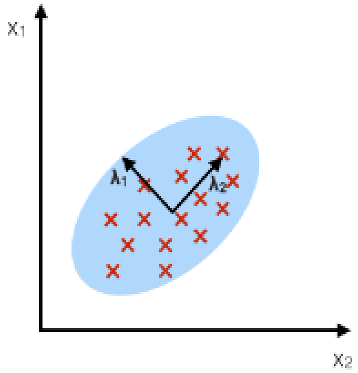


Source: Hands-On Machine Learning with R

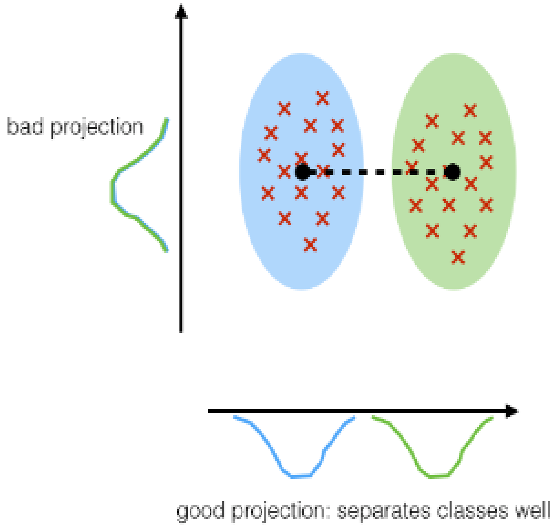
3. Machine Learning Algorithms

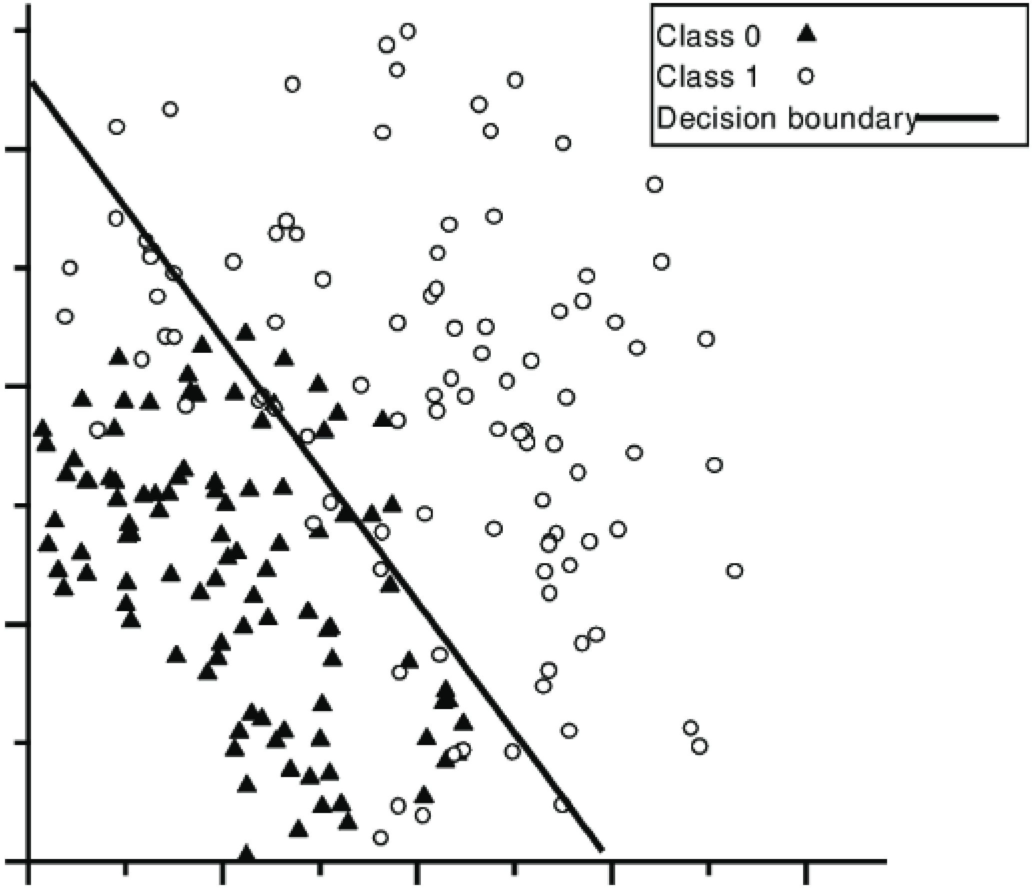
2. Various Machine Learning Algorithms
4) Discriminant Analysis

PCA:
component axes that maximize the variance

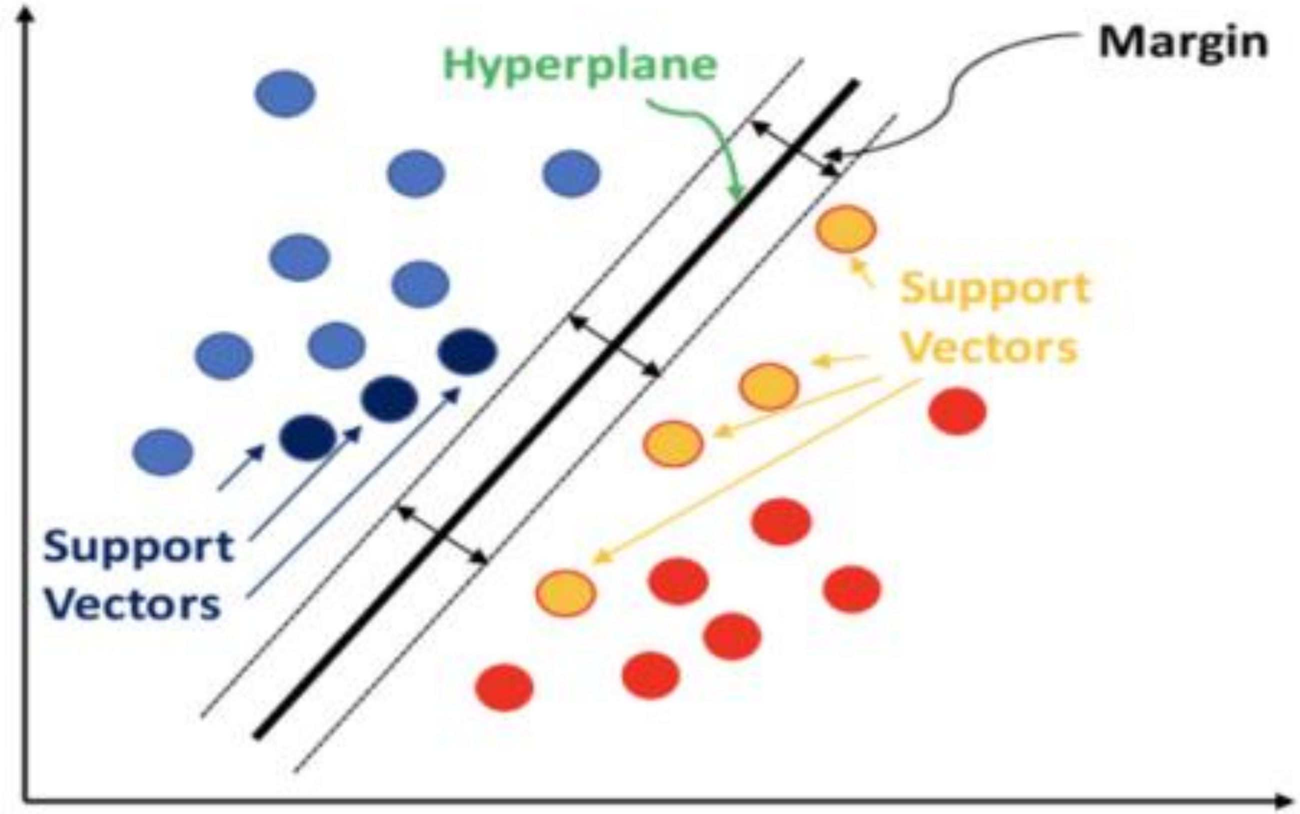


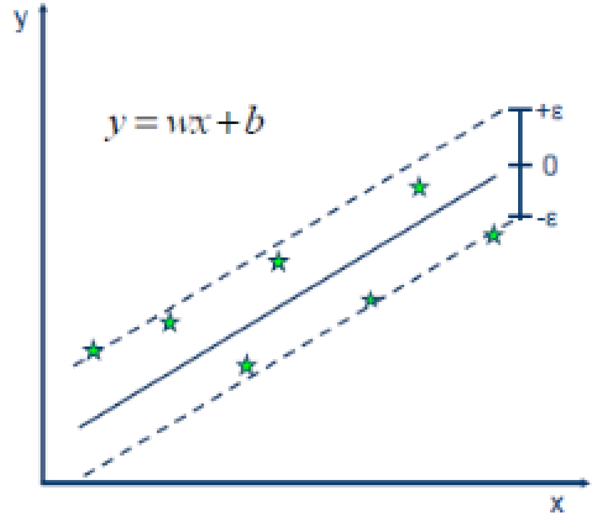
LDA:
maximizing the component axes for class-separation





2. Various Machine Learning Algorithms
5) Support Vector Machine



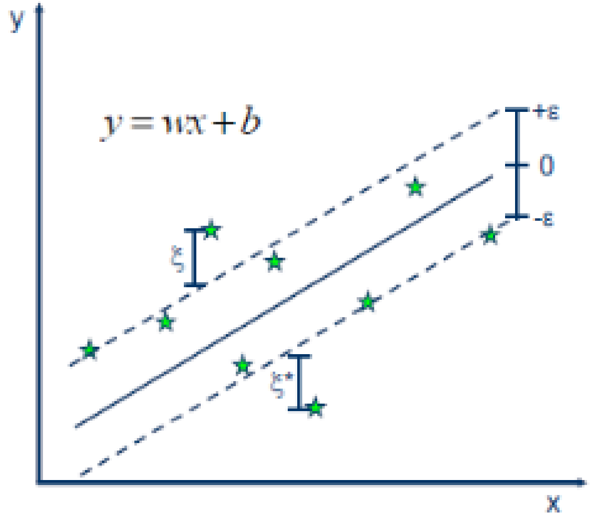


- Solution:

$$\min \frac{1}{2}\|w\|^2$$

- Constraints:

$$y_i - wx_i - b \le \varepsilon$$
$$wx_i + b - y_i \le \varepsilon$$



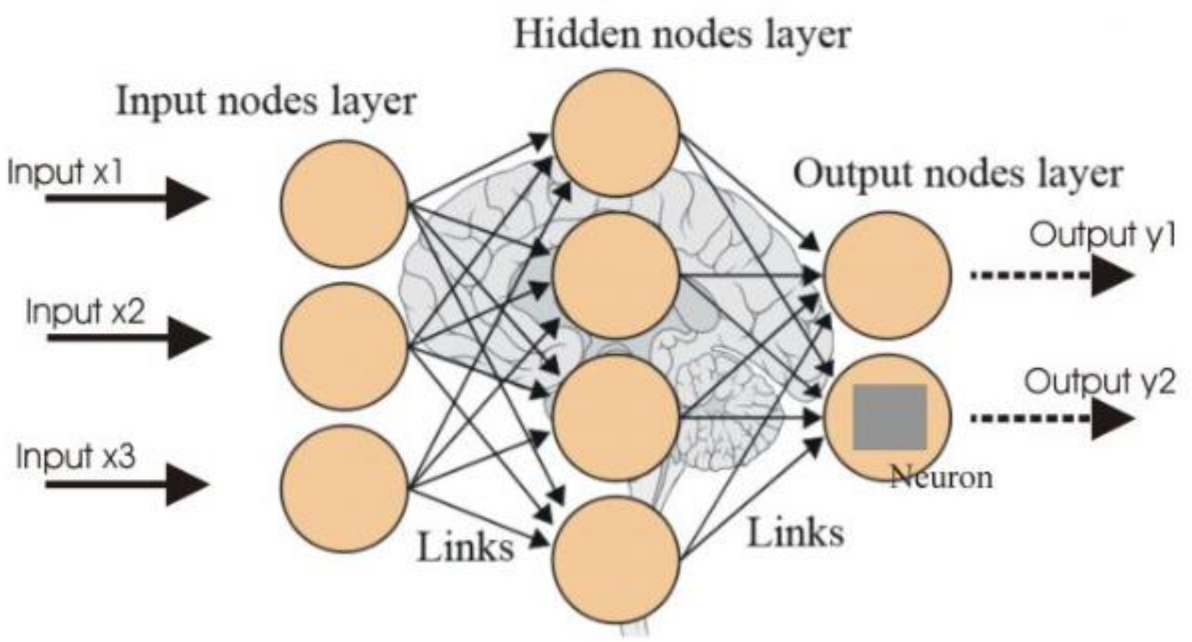
- Minimize:

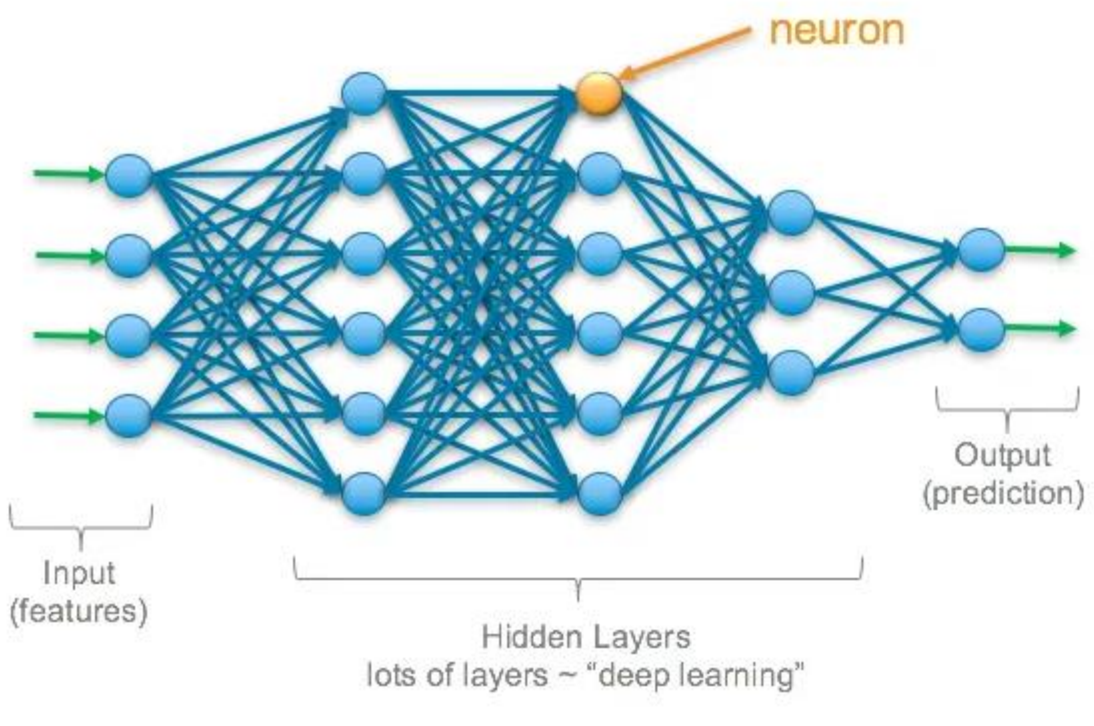
$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

- Constraints:

$$y_i - wx_i - b \le \varepsilon + \xi_i$$
$$wx_i + b - y_i \le \varepsilon + \xi_i^*$$
$$\xi_i, \xi_i^* \ge 0$$

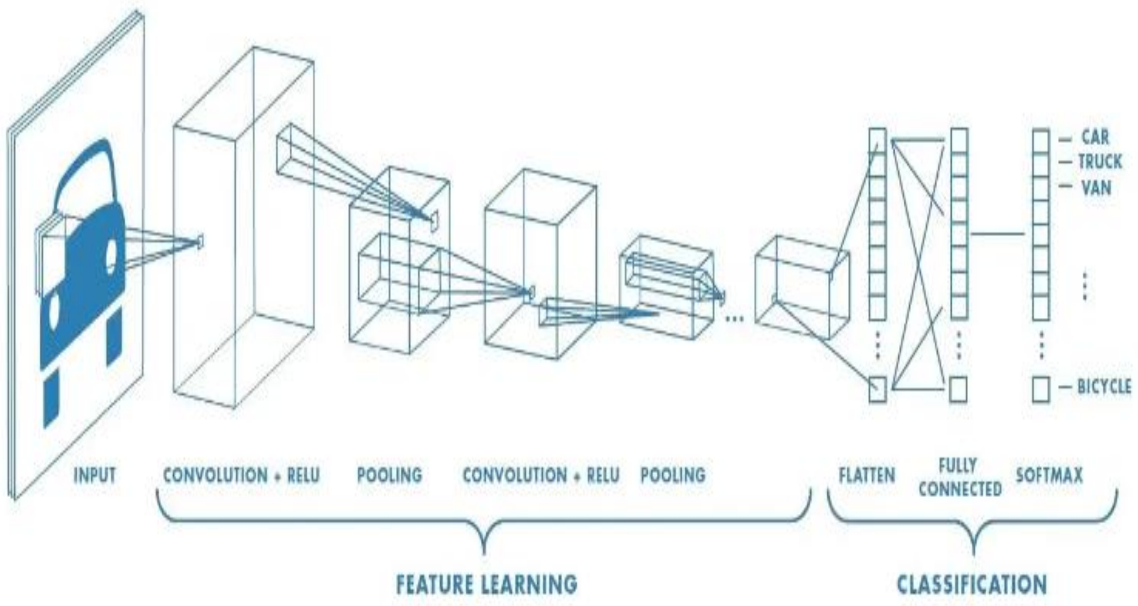
2. Various Machine Learning Algorithms
6) Deep Neural Network

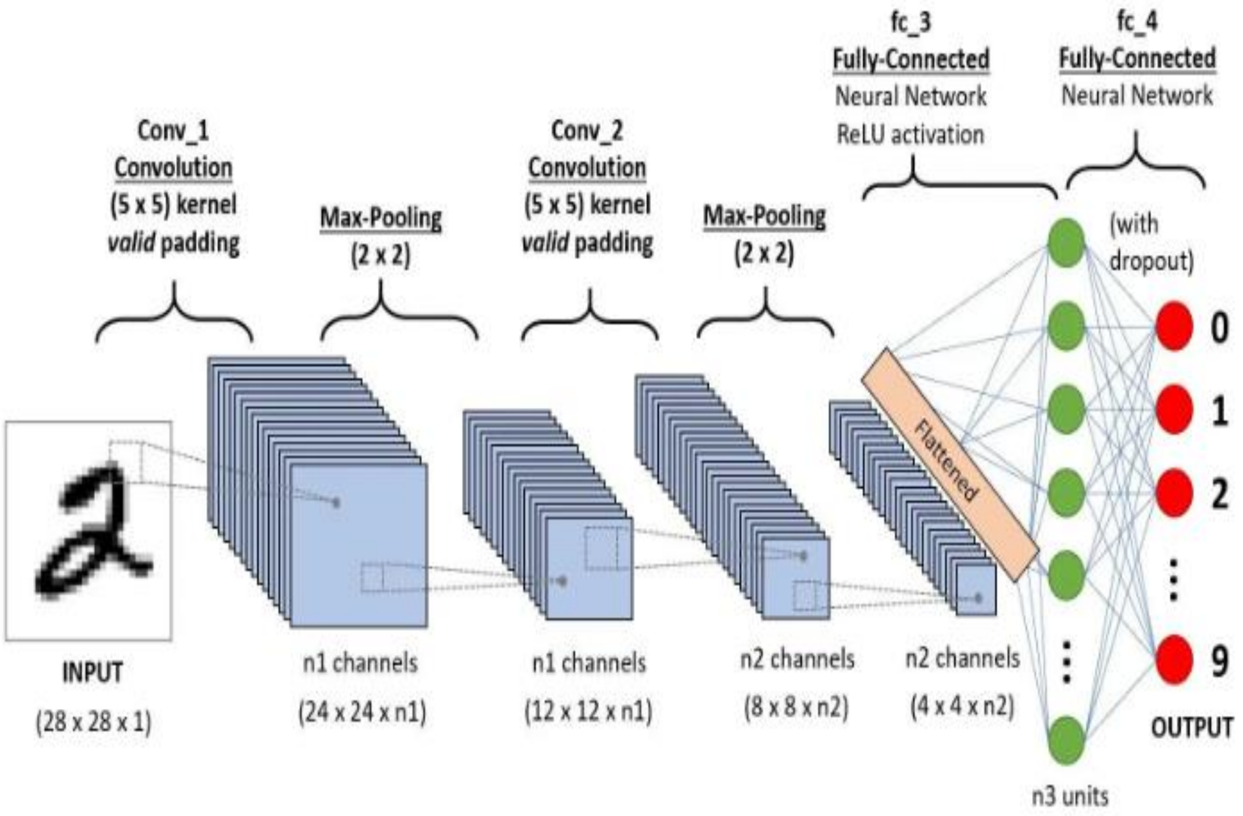




2. Various Machine Learning Algorithms
7) Convolutional Neural Network

By Sumit Saha | Saturday, December 15, 2018 | Data Science & ML





감사합니다