

## Penerapan Metode Naïve Bayes Untuk Mengetahui Kualitas Air Di Jakarta

Yunita Sartika Sari

Fakultas Ilmu Komputer, Universitas Mercu Buana  
Jl. Meruya Selatan No. 1, Kembangan, Jakarta Barat, 11650  
yunita.sartika@mercubuana.ac.id

### ABSTRACT

*Water is a chemical compound that is needed for the survival of living things on earth. The largest area on planet earth is water which covers almost 71% of the area on earth. Water is also a very important substance on earth that is needed by all living things ranging from plants, animals and humans. It is necessary to monitor and treat the environment around the water source so that it can produce clean water quality. In this study, the Naïve Bayes algorithm is applied to determine water quality in the Jakarta area and produces an accuracy rate of water classification results in the DKI Jakarta area of 50.6%.*

**Keywords:** *nave bayes, quality, water, Jakarta*

### ABSTRAK

Air merupakan senyawa kimia yang sangat dibutuhkan bagi kelangsungan hidup makhluk hidup yang ada di bumi. Wilayah terluas di planet bumi merupakan air yang menutupi hampir 71% wilayah yang ada di bumi. Air juga merupakan zat yang sangat penting yang ada di bumi yang sangat dibutuhkan oleh semua makhluk hidup mulai dari tumbuhan, hewan dan manusia. Dibutuhkan pengawasan dan pengolahan lingkungan sekitar sumber air sehingga dapat menghasilkan kualitas air yang bersih. Pada penelitian ini menerapkan algoritma Naïve Bayes untuk mengetahui kualitas air di wilayah Jakarta dan menghasilkan tingkat akurasi dari hasil klasifikasi air di wilayah DKI Jakarta sebesar 50,6% .

**Kata kunci :** *naïve bayes, kualitas, air, Jakarta*

## I. PENDAHULUAN

Air merupakan sumberdaya alam yang mempunyai fungsi sangat penting bagi kehidupan manusia dan makhluk hidup lainnya serta sebagai modal dasar dalam pembangunan. Dengan perannya yang sangat penting, air akan mempengaruhi dan dipengaruhi oleh kondisi/komponen lainnya.

Namun tidak semua zat dan mineral yang terkandung di air dapat dicerna dan diterima dengan baik oleh tubuh manusia. Air juga rentan terkontaminasi oleh bakteri-bakteri dan zat mineral yang berbahaya bagi tubuh manusia. Hal tersebut bisa terjadi dikarenakan tercemarnya sumber air atau tercemarnya lingkungan di sekitar sumber air. Begitu pentingnya peranan air dalam kehidupan sehingga dapat dinyatakan bahwa kualitas air dapat digunakan sebagai indikator tingkat kesehatan manusia (Situmorang, 2017).

Dibutuhkan pengawasan dan pengolahan lingkungan sekitar sumber air sehingga dapat menghasilkan kualitas air yang bersih sesuai dengan standar kualitas air bersih dan memenuhi standar air yang layak dikonsumsi oleh manusia. Untuk menentukan klasifikasi kualitas air bersih terdapat banyak metode klasifikasi yang dapat digunakan, salah satunya dengan menggunakan algoritma naïve bayes. Dalam penelitian sebelumnya oleh Rifwan Hamidi dengan menggunakan metode *Learning Vector Quantization* untuk klasifikasi kualitas air sungai menggunakan parameter sejumlah 7 masukan dengan menghasilkan akurasi sebesar 81.13%. Penelitian lain yang dilakukan oleh Mila Listiana pada tahun 2015 dalam kasus identifikasi tumbuh kembang anak balita dengan perbandingan algoritma *Decision Tree(C4.5)* dan *Naïve Bayes*, diperoleh hasil pengujian rata-rata nilai akurasi metode Naïve Bayes sebesar 96,89% dan algoritma *Decision Tree(C4.5)* sebesar 89,78%. Penelitian lain yang dilakukan oleh Khafiizh Hastuti pada tahun 2012 pada prediksi data mahasiswa nonaktif dengan melakukan perbandingan metode klasifikasi *Logistic Regression*, *Decision Tree*, *Neural Network* dan *Naïve Bayes*. Penelitian tersebut menghasilkan tingkat akurasi metode *Logistic Regression* sebesar 81,64%, metode *Decision Tree* sebesar 95,29%, metode *Neural*

Network sebesar 94,56% dan metode *Naïve Bayes* sebesar 93,47%. Selain itu terdapat penelitian lain yang dilakukan oleh Santoso pada tahun 2016 dengan melakukan komparasi metode *K-Nearest Neighbor* dan *Learning Vector Quantization* (LVQ) dengan studi kasus peramalan klasifikasi tingkat kemiskinan menghasilkan akurasi metode *K-Nearest Neighbor* sebesar 93.52% dan *Learning Vector Quantization* (LVQ) sebesar 75.93%. Pada penelitian sebelumnya terdapat perbedaan tingkat akurasi beberapa metode untuk klasifikasi.

Berdasarkan dari beberapa penelitian sebelumnya, *Naïve Bayes* merupakan metode yang memiliki akurasi cukup tinggi. Untuk itu penelitian ini, peneliti akan menerapkan algoritma *Naïve Bayes* untuk mengklasifikasikan kualitas air bersih. Berdasarkan uraian latar belakang, maka judul yang diambil dalam skripsi ini adalah “Penerapan Metode *Naïve Bayes* Untuk Mengetahui Kualitas Air di Jakarta”.

## II. LANDASAN TEORI

### A. Pengertian Metode Classification

Metode *Classification* adalah sebuah metode dari data mining yang digunakan untuk memprediksi kategori atau kelas dari suatu data instance berdasarkan sekumpulan atribut-atribut dari data tersebut.

Terdapat beberapa algoritma data mining yang menggunakan metode *Classification* ini, seperti C4.5, CMAR, *Naïve Bayes*, *K Nearest Neighbor* dan algoritma yang penulis implemetasikan, CART. (Han, 2016)

### B. Pengertian Naïve Bayes

*Naïve Bayes Classifier* merupakan sebuah metoda klasifikasi yang berakar pada teorema Bayes. Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari *Naïve Bayes Classifier* ini adalah asumsi yang sangat kuat (naïf) akan independensi dari masing-masing kondisi / kejadian.

Menurut Olson Delen (2008) menjelaskan *Naïve Bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi obyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari “master” tabel keputusan.

*Naïve Bayes Classifier* bekerja sangat baik dibanding dengan model *classifier* lainnya. Hal ini dibuktikan oleh Xhemali, Hinde Stone dalam jurnalnya “*Naïve Bayes vs. Decision Trees vs. Neural Networks in the Classification of Training Web Pages*” mengatakan bahwa “*Naïve Bayes Classifier* memiliki tingkat akurasi yg lebih baik disbanding model *classifier* lainnya”.

Keuntungan penggunaan adalah bahwa metoda ini hanya membutuhkan jumlah data pelatihan (*training data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. Karena yang diasumsikan sebagai *variable independent*, maka hanya varians dari suatu variable dalam sebuah kelas yang dibutuhkan untuk menentukan klasifikasi, bukan keseluruhan dari matriks kovarians.

## III. METODE DAN LANGKAH PENELITIAN

Metode yang digunakan pada penelitian ini adalah metode klasifikasi menggunakan algoritma *Naïve Bayes*. Penelitian ini bertujuan untuk mengklasifikasikan kualitas air di wilayah DKI Jakarta. Adapun klasifikasi kualitas air terdiri dari kondisi baik, tercemar ringan, tercemar sedang, dan tercemar berat.

Tahapan penelitian di atas yang digunakan dalam penelitian ini yang dapat dijelaskan sebagai berikut :

### Tahap 1 : Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data untuk mengetahui variable yang digunakan. Metode pengumpulan data yang digunakan yaitu, wawancara, literatur pustaka, dan observasi.

### Tahap 2: Analisa Masalah

Pada tahap ini akan melakukan analisa masalah untuk menerapkan metode yang dipilih yaitu *Naïve Bayes*. Pada tahap ini, peneliti dapat menentukan persepsi dan variable apa saja yang akan diperoleh beserta visualisasinya.

### Tahap 3: Penerapan Metode

Tahap ini merupakan tahap penerapan metode dari hasil analisa masalah yang pada proses pengumpulan data.

#### IV. HASIL DAN PEMBAHASAN

##### A. Parameter Kualitas Air

Indikator yang sering kali digunakan dalam pemeriksaan pencemaran air adalah pH, oksigen terlarut (*Dissolved Oxygen*, DO), kebutuhan oksigen biokimia (*Biochemiyca Oxygen Demand*, BOD) serta kebutuhan oksigen kimiawi (*Chemical Oxygen Demand*, COD). Namun pada pengujian ini menggunakan beberapa parameter yang terdapat pada sensor air yaitu *potential of Hydrogen* (pH), suhu atau *Temperature* dan *Total Dissolved Solids* (TDS).

Pada Tabel 1: merupakan nilai berdasarkan standar kualitas air yang dipakai dengan tingkat kondisi kualitas air. Pada range pH 6.5 – 8.5 menunjukkan kondisi alami airbelum tercemar, range pH kurang dari 6 menunjukkan kondisi asam, dan range pH lebih dari 8 menunjukkan kondisi basa. Pada range suhu 26°-32° menunjukkan kondisi netral, range suhu kurang dari 26° menunjukkan kondisi air dingin, dan range suhu lebih dari 32° menunjukkan kondisi air panas. Pada range TDS 0 – 250 menunjukkan kondisi baik, range TDS 250 - 400 menunjukkan kondisi cukup dan range TDS lebih dari 400 menunjukkan kondisi buruk.

Tabel 1. Nilai Standar Parameter Kualitas Air

Parameter Kualitas Air		
Nama Parameter	Nilai Range pH	Kondisi
pH	6.5 - 8.5	Kondisi Alami
	<6	Kondisi Asam
	>8	Kondisi Basa
Suhu	26 - 32	Netral
	<26	Dingin
	>32	Panas
TDS	0-250	Baik
	250-400	Cukup
	>400	Buruk

##### B. Model Klasifikasi yang Digunakan

*Naive Bayes Classifier* merupakan salah satu metode di dalam data mining untuk mengklasifikasikan data. Cara kerja dari metode *Naive Bayes Classifier* menggunakan perhitungan probabilitas. Konsep dasar yang digunakan oleh *Naive bayes* adalah teorema Bayes, yaitu teorema yang digunakan dalam statistika untuk menghitung suatu peluang, *Naive Bayes Classifier* menghitung peluang dari satu kelas dari masing- masing kelompok atribut yang ada dan menentukan kelas yang paling optimal. Proses pengelompokan atau klasifikasi dibagi menjadi dua fase yaitu *learning/training* dan *testing/classify*. Pada fase learning, sebagian data yang telah diketahui kelas, datanya diumpankan untuk membentuk model perkiraan. Kemudian pada fase testing, model yang sudah terbentuk diuji dengan sebagian data.

Dasar dari Naive Bayes yang dipakai dalam pemrograman adalah rumus Bayes:

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan:

- X: Data dengan class yang belum diketahui
- H: Hipotesis data merupakan suatu class spesifik
- P(H|X): Probabilitas hipotesis H berdasarkan kondisi X (*posteriori probabilitas*)
- P(H): Probabilitas hipotesis H (*Prior probabilitas*)
- P(X|H): Probabilitas X berdasarkan kondisi pada hipotesis H
- P(X): Probabilitas X

Tabel 2. Data Training

No	pH	Suhu	TDS	Kondisi Kelas
1	7	31	95	Kondisi Baik
2	6	29	127	Kondisi Baik
...	...	...	...	Kondisi Baik
250	8	30	130	Kondisi Baik
251	8	27	259	Tercemar Ringan
252	7	31	264	Tercemar Ringan
...	...	...	...	Tercemar Ringan
500	6	28	384	Tercemar Ringan
501	28	33	300	Tercemar Ringan
502	8	31	783	Tercemar Ringan
...	...	...	...	Tercemar Ringan
750	5	30	363	Tercemar Ringan
751	4	30	949	Tercemar Berat
752	11	28	461	Tercemar Berat
...	...	...	...	Tercemar Berat
1000	5	18	496	Tercemar Berat

Berikut merupakan penamaan label untuk nilai parameter kualitas air yang digunakan, terdapat tiga parameter kualitas air dan label setiap masing-masing nilai parameter yang digunakan:

- Nilai pH

- a) 0 = Kondisi netral, range nilai = 6 - 8
- b) 1 = Kondisi asam, range nilai = 0 - 5
- c) 1 = Kondisi basa, range nilai = 9 - 14

- Nilai suhu

- a) 0 = Kondisi baik, range nilai = 26° - 32°
- b) 1 = Kondisi dingin, range nilai = < 26°
- c) 1 = Kondisi panas, range nilai = > 32°

- Nilai TDS

- a) 0 = Kondisi baik, range nilai = 0 - 250
- b) 1 = Kondisi cukup, range nilai = 250 - 400
- c) 2 = Kondisi buruk, range nilai = > 400

Berikut merupakan penamaan untuk label kelas kategori kondisi air, terdapat empat label dan kelas kategori kondisi air yang digunakan yaitu kondisi baik, tercemar ringan, tercemar sedang dan tercemar berat.

- 1) 0 = Kondisi baik
- 2) 1 = Tercemar ringan
- 3) 2 = Tercemar sedang
- 4) 3 = Tercemar berat

Tabel 3. *Rules Naïve Bayes*

Label pH	Label Suhu	Label TDS	Label Kelas
0	0	0	0
0	0	1	1
0	0	2	2
0	1	0	1
0	1	1	2
0	1	2	3
1	0	0	1
1	0	1	2
1	0	2	3
1	1	0	2
1	1	1	2
1	1	2	3

Adapun nilai probabilitas dengan kategori kondisi air dapat dilihat pada Tabel 4 dan nilai probabilitas setiap parameter dapat dilihat pada Tabel 5.5 nilai pH, Tabel 5.6 nilai suhu, dan Tabel 5.7 nilai TDS berikut.

Tabwl 4. Probabilitas Kategori Kelas Kondisi Air

Jumlah Kejadian "Kondisi Air"				Probabilitas			
Baik	Ringan	Sedang	Buruk	Baik	Ringan	Sedang	Buruk
250	250	250	250	0,25	0,25	0,25	0,25

Tabel 5. Probabilitas Nilai PH

Keadaan pH	Jumlah Kejadian "Kondisi Air"				Probabilitas			
	Baik	Ringan	Sedang	Buruk	Baik	Ringan	Sedang	Buruk
Netral	250	166	100	83	0,996	0,6667	0,2024	0,3333
Asam dan Basa	0	84	150	167	0,004	0,3333	0,7976	0,6667
Jumlah	250	250	250	250	1	1	1	1

Pada Tabel 5: merupakan perhitungan probabilitas nilai pH. Dapat dilihat bahwa Keadaan pH netral dalam kondisi baik memiliki probabilitas 0.996, pH netral kondisi ringan memiliki probabilitas 0.6667, pH netral kondisi sedang memiliki probabilitas 0.2024, pH netral kondisi buruk 0.3333. Keadaan pH asam dan basa dalam kondisi baik memiliki probabilitas 0.004, pH asam dan basa kondisi ringan memiliki probabilitas 0.3333, pH asam dan basa kondisi sedang memiliki probabilitas 0.7976, pH asam dan basa kondisi buruk 0.6667.

Tabel 6. Probabilitas Nilai Suhu

Keadaan Suhu	Jumlah Kejadian "Kondisi Air"				Probabilitas			
	Baik	Ringan	Sedang	Buruk	Baik	Ringan	Sedang	Buruk
Baik Dingin dan Panas	250	167	100	83	0,996	0,6667	0,3532	0,3333
	0	83	150	167	0,004	0,3333	0,6468	0,6667
Jumlah	250	250	250	250	1	1	1	1

Pada Tabel 6: merupakan perhitungan probabilitas nilai suhu. Dapat dilihat bahwa Keadaan suhu baik dalam kondisi baik memiliki probabilitas 0.996, suhu baik kondisi ringan memiliki probabilitas 0.6667, suhu baik kondisi sedang memiliki probabilitas 0.3532, suhu baik kondisi buruk 0.3333.

Tabel 7. Probabilitas Nilasi TDS

Keadaan TDS	Jumlah Kejadian "Kondisi Air"				Probabilitas			
	Baik	Ringan	Sedang	Buruk	Baik	Ringan	Sedang	Buruk
Baik	250	167	50	0	0,9921	0,664	0,2016	0,0039
Cukup	0	83	150	0	0,0039	0,332	0,5969	0,0039
Buruk	0	0	50	250	0,004	0,004	0,2015	0,9921
Jumlah	250	250	250	250	1	1	1	1

Pada tabel 7: merupakan probabilitas nilai TDS. Dapat dilihat bahwa Keadaan TDS baik dalam kondisi baik memiliki probabilitas 0.9921, TDS baik kondisi ringan memiliki probabilitas 0.664, TDS baik kondisi sedang memiliki probabilitas 0.2016, TDS baik kondisi buruk 0.0039. Keadaan TDS cukup kondisi baik memiliki probabilitas 0.0039, TDS cukup kondisi ringan memiliki probabilitas 0.332, TDS cukup kondisi sedang memiliki probabilitas 0.5969, TDS cukup kondisi buruk 0.0039. Keadaan TDS buruk kondisi baik memiliki probabilitas 0.004, TDS buruk kondisi ringan memiliki probabilitas 0.004, TDS buruk kondisi sedang memiliki probabilitas 0.2015, TDS buruk kondisi buruk 0.9921.

### C. Pengukuran Tingkat Akurasi

Dalam melakukan klasifikasi diharapkan dapat melakukan klasifikasi dengan benar, sehingga dalam klasifikasi harus diukur tingkat akurasi. Akurasi didapatkan dari jumlah data yang diprediksi benar dibagi dengan jumlah data yang diprediksi. Berikut merupakan formula untuk menghitung tingkat akurasi menggunakan persamaan.

$$Akurasi = \frac{\sum i}{\sum n} * 100\%$$

Keterangan:

- $\sum i$  = Jumlah data yang diprediksi benar
- $\sum n$  = Jumlah data yang diprediksi

Dalam melakukan prediksi klasifikasi, dapat melakukan klasifikasi pada semua objek dengan benar. Percobaan dilakukan dengan menggunakan 506 data testing atau uji dan 1000 data training yang memberikan hasil akurasi sebesar 50,6%. Untuk hasil akurasi yang baik dalam klasifikasi sistem ini dengan melakukan perhitungan rules Naive Bayes yang tepat dan jumlah data training yang lebih banyak dibandingkan dengan jumlah data testing pengujian.

## V. KESIMPULAN

Dari hasil klasifikasi menggunakan naïve bayes, tingkat akurasi dari hasil klasifikasi air di wilayah DKI Jakarta sebesar 50,6% .

## DAFTAR PUSTAKA

- Alfa Saleh. 2015. Implementasi metode klasifikasi naïve bayes dalam memprediksi besarnya penggunaan listrik rumah tangga. *Creative Information Technology Journal*, 2(3):207–217
- Ali Masduqi, Eddy S Soedjono, Noor Endah, and Wahyono Hadi. 2009. Prediction of rural water supply systems sustainability using a mathematical model. *Jurnal Purifikasi*, 10(2):155–164
- Alieja Muhammad Putrada, Maman Abdurrohman, and Aji Gautama Putrada. 2018. Increasing smoke classifier accuracy using naïve bayes method on internet of things. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 4(1):19–26
- Budi Santosa. Data mining teknik pemanfaatan data untuk keperluan bisnis. Yogyakarta: Graha Ilmu, 978(979):756, 2007.
- Burhan Alfironi Mukhtar, Noor Akhmad Setiawan, and Teguh Bharata Adji. 2015. Analisis perbandingan tingkat akurasi algoritma naïve bayes classifier dengan correlated-naïve bayes classifier. *SEMNAS TEKNOMEDIA ONLINE*, 3(1):2–1
- Chongsheng Zhang, Changchang Liu, Xiangliang Zhang, and George Alpanidis. 2017. An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications*, 82:128–150
- Daniela Xhemali, Chris J Hinde, and Roger G Stone. 2009. Naïve bayes vs. decision trees vs. neural networks in the classification of training web pages

- Dewan Md Farid and Mohammad Zahidur Rahman. 2010. Anomaly network intrusion detection based on improved self adaptive bayesian algorithm. *JCP*, 5(1):23–31
- Hartatik Hartatik, Andri Syafrianto, and Wiwi Widayani. 2018. Perbandingan klasifikasi pencemaran air sungai dengan metode backpropagation dan naïve bayes. *Data Manajemen dan Teknologi Informasi (DASI)*, 18(4):67–71
- Hefni Effendi. 2003. Telaah kualitas air, bagi pengelolaan sumber daya dan lingkungan perairan. Kanisius
- Kusnawi. 2007. “Pengantar Solusi Data Mining”. STMIK AMIKOM, Yogyakarta.
- Kustiyahningsih Yeni, dkk. 2013. “Sistem Pendukung Keputusan untuk Menentukan Jurusan pada Siswa SMA Menggunakan Metode KNN dan SMART”. Universitas Trunojoyo, Madura.
- Kusumadewi, Sri. 2009. “Klasifikasi Status Gizi Menggunakan Naive Bayesian Classification”. Universitas Islam Indonesia, Yogyakarta.
- Lestiana, Mila. 2015. “Perbandingan Algoritma Decision Tree (C4.5) dan Naive Bayes pada Data Mining untuk Identifikasi Tumbuh Kembang Anak Balita”. Universitas Muhammadiyah, Surakarta. Peraturan Menteri Republik Indonesia nomor 492 tahun 2010 tentang Persyaratan Kualitas Air minum. Jakarta : Kementerian Kesehatan Republik Indonesia. Peraturan Pemerintah Republik Indonesia nomor 82 tahun 2001 tentang Pengelolaan Kualitas Air dan Pengendalian Pencemaran Air. Jakarta.
- Mohamad Sakizadeh. 2015. Assessment the performance of classification methods in water quality studies, a case study in karaj river. *Environmental monitoring and assessment*, 187(9):573
- Pablo Bermejo, José A Gámez, and José M Puerta. 2014. Speeding up in incremental wrapper feature subset selection with naïve bayes classifier. *Knowledge-Based Systems*
- Riyadh Arridha. 2018. Design and implementation of iot-big data analytic for smart environment monitoring system.
- Santoso. 2016. “Perbandingan Metode KNearest Neighbor(K-NN) dan Learning Vector Quantization (LVQ) untuk Permasalahan Klasifikasi Tingkat Kemiskinan”. Institut Teknologi Sepuluh November, Surabaya.
- Situmorang, Manihar. 2017. “Kimia Lingkungan”. Depok: PT RajaGrafindo
- Tri Vuldari, Retno. 2017. “Data Mining Teori dan Aplikasi Rapidminer”. Surakarta.
- Saeful Bachrein. 2015. Pengembangan daerah aliran sungai (das) cikapundung: Diagnostik wilayah. *Jurnal Bina Praja: Journal of Home Affairs Governance*, 4(4):227–236
- Sritrusta Sukaridhoto, Dadet Pramadihanto, Muhammad Alif, Andrie Yuwono, Nobuo Funabiki, et al. 2015. A design of radio- controlled submarine modification for river water quality monitoring. In 2015 International Seminar on Intelligent Technology and Its Applications (ISITIA), pages 75–80. IEEE
- Sritrusta Sukaridhoto, Rahardhita Widyatra Sudiby, Widi Sarinastiti, Rizky Dharmawan, Atit Sasono, Ahmad Andika Saputra, and Shiori Sasaki. 2016. Design and development of a portable low-cost cots-based water quality monitoring system. In 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), pages 635–640. IEEE
- Shumei Wang, Zhaoji Zhang, Zhilong Ye, Xiaojun Wang, Xiangyu Lin, and Shaohua Chen. 2013. Application of environmental internet of things on water quality management of urban scenic river. *International Journal of Sustainable Development & World Ecology*, 20(3):216–222
- Xindong Wu and Vipin Kumar. 2009. *The top ten algorithms in data mining*. CRC press