

Simulasi Prakiraan Dan Klasifikasi Hujan Wilayah Kota Jakarta dengan Metode Decision Tree

Arif Ega Prakosa¹, Ahmad Fawaid², Irkhas Nusantara³, Faisal Amri⁴, Aries Saifudin⁵
Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email : ¹arifega12@gmail.com, ²fawaidahmad946@gmail.com, ³irkhasan2@gmail.com,
⁴faisalamryy@gmail.com, ⁵aries.saifudin@unpam.ac.id

Submitted Date: November 18th, 2020
Revised Date: November 25th, 2020

Reviewed Date: November 24th, 2020
Accepted Date: April 01st, 2021

Abstract

Rain is a natural occurrence that occurs in the hydrological and climatic circulation. Based on the ups and downs of the hydrological circulation, one of the sources of water is rain. rain is very useful in life, because rain can meet the needs of water for living creatures. However, rain can also cause floods. A flood tragedy can cause loss and casualties. Then our activities are checking the forecast simulation forecast and rain classification using the accurate decision tree method. We take forecast and classification data because the weather in DKI Jakarta is currently very difficult to predict. Next, we use a decision tree for data mining with a dataset spanning 5 years from 2011-2015. The problem here is the weather in DKI Jakarta which is very difficult to predict.

Keyword: Weather Data; Data Mining; Decision tree

Abstrak

Hujan adalah sebuah kejadian alam yang terdapat dalam siklus hidrologi dan iklim. Berdasarkan naik turun nya sirkulasi hidrologi salah satu sumber air adalah hujan. hujan sangat bermanfaat dalam kehidupan, karena hujan dapat memenuhi kebutuhan air untuk makhluk hidup. tetapi, hujan juga dapat menimbulkan bencana banjir. Tragedi banjir dapat menyebabkan kerugian dan korban jiwa. Kemudian kegiatan yang kami lakukan adalah meneliti untuk mengetahui simulasi prakiraan dan klasifikasi hujan dengan menggunakan metode Decision Tree seberapa tinggi intensitas akuratnya. Kami mengambil data prakiraan dan klasifikasi karena cuaca di DKI Jakarta saat ini sangat susah untuk diprediksi. Selanjutnya kita menggunakan decision tree untuk data miningnya dengan dataset jangkauan 5 tahun dari 2011-2015. Masalah di sini yaitu cuaca di DKI Jakarta yang sangat sulit diprediksi.

Kata kunci: Data Cuaca; Data Mining; Decision tree

1 Pendahuluan

Air adalah sesuatu yang sangat penting bagi kehidupan makhluk hidup di dunia ini. Jadi air merupakan suatu hal yang sangat berharga. Air juga dapat dimanfaatkan untuk keperluan di berbagai bidang, misalnya untuk kebutuhan sehari-hari dan PLTA (Pembangkit Listrik Tenaga Air). Bisa disebut air merupakan anugerah dari tuhan yang membawa kesejahteraan bagi manusia dan semua makhluk hidup .

Dari dinamika siklus hidrologi salah satu sumber air utama adalah hujan. Dari alam hujan terjadi dari proses menguapnya air di udara yang selanjutnya membentuk gumpalan awan. Sifat hujan tergantung dengan kondisi cuaca yang

terjadi pada saat itu. Ketersediaan air dari alam dalam jangkauan global yaitu tetap, hanya terjadi variasi baik pada ruang ataupun waktu pada saat skala regional. Sehingga dalam suatu pengamatan terhadap cuaca bisa diprediksi oleh cara simulasi (T. & Sari, 2017). Untuk mengetahui simulasi prakiraan dan klasifikasi cuaca ini digunakan hasil data dari tahun ketahun pada setiap bulannya untuk menjadikan acuan cuaca pada bulan-bulan simulasinya.

Membuat materi simulasi prakiraan cuaca karena akhir – akhir ini cuaca susah ditebak tidak sesuai dengan bulannya sehingga diambil data untuk prakiraan ditarik dari 5 tahun sebelumnya bagaimana data hujan tapi hanya 50% akurat

prakiraannya (Siregar, Faisal, Cahyana, & Priyatna, 2020). Tingkat hujan dapat diukur dengan menggunakan alat pengukur hujan, alat yang digunakan untuk mengukur tingkat besar kecilnya hujan. Alat ini bekerja untuk mengumpulkan data curah hujan pada setiap tahun dan bulannya.

Kemudian Proses prakiraan cuaca memerlukan banyak data cuaca, jumlah data yang besar (Dhawangkharu & Riksakomara, 2017). Data yang besar dapat mempercepat dan tepat dalam melakukan proses ini dengan menggunakan metode decision tree.

Dalam pengambilan data yang akan digunakan pada topic yang kami pakai dari data internet dari berbagai website resmi dari tahun ketahun untuk menentukan data yang akan digunakan (Lumbansiantar, 2019). Decision tree yang digunakan untuk mensimulasikan dan mengklasifikasikan cuaca akan memudahkan mendapatkan hasil yang dibutuhkan (Mendrofa, 2019).

Decision tree digunakan bermanfaat untuk membreakdown proses pengambilan keputusan yang susah dipahami agar mudah dipahami (Rosdiana & Rismayana, 2018). Pengambilan keputusan akan lebih mengutamakan interpretasi solusi dari sebuah permasalahan.

Data mining untuk memprakiraan cuaca dengan menggunakan data cuaca beberapa tahun ke belakang (Maesaroh & Kusri, 2017). Bisa dibidang data mining berfungsi mendeskripsikan apa yang akan terjadi pada kedepannya. Salah satu contoh teknik descriptive data mining dan predictive data mining adalah klustering dan klasifikasi (Saifudin, 2018). Klustering digunakan pada saat data tidak diketahui bagaimana dikelompokkan atau pengelompokan data dengan cara membagi semua data menjadi kelompok-kelompok yang memiliki kesamaan (Gunadi & Dewi, 2018). Setelahnya, suatu sistem akan mengklarifikasi data yang baru kedalam kelompok yang ada dan tidak ada tambahan kelompok data baru.

2 Metodologi

2.1 Metode

Pengambilan data

Pengambilan sumber tersebut dapat berupa buku bacaan, dan tesis yang berhubungan dengan topik yang diambil, maupun referensi

langsung dari internet. Penelitian ini menggunakan pengambilan yang langsung dari internet.

Pengolahan data awal

Selanjutnya dalam pengolahan data berisikan pengecekan data yang mentah tadi yaitu melengkapkan data yang ada itu dengan menghapus data yang tidak digunakan supaya lebih mudah untuk mengelompokkan datanya tersebut (Mujiasih, 2011).

Analisis Kebutuhan dan Perancangan Sistem

Selanjutnya pada tahap analisis kebutuhan sistem yang akan dibuat, kemudian membuat struktur desain sistem yang telah dianalisa.

Implementasi dan Pengujian Sistem

Sistem dibangun sesuai rancangan yang telah disetujui bersama, dalam hal ini sistem dibangun dengan menggunakan aplikasi berbentuk PHP dan Database. Selanjutnya menguji sistem yang terbentuk.

Evaluasi dan validasi hasil

Mengevaluasi dilakukan dengan mengamati hasil menggunakan algoritmanya. Validasi digunakan untuk memastikan bahwa hasil yang terakhir sudah sesuai dengan keinginan awal perancangan yang dibikannya. Pengukuran kinerja dilakukan dengan membandingkan nilai ketepatan prediksi dan kesempurnaan data yang didapat (Laila & Setyawan, 2020).

Desain Rancangan

Mengambil desain rancangan untuk memasukkan data setahun sebelumnya guna mengetahui prakiraan cuaca dan klasifikasi cuaca yang akan datang.

2.2 Data

Dataset BMKG

Ini data yang diambil dari dataset bmkg yang berisi tingkat jumlah curah hujan dan jumlah hari hujan di semua provinsi Indonesia dan yang digunakan hanya data provinsi DKI Jakarta dari tahun 2011 hingga 2015.

Tabel 1 Data BMKG Indonesia

Tahun		2011		2012		2013		2014		2015	
Provinsi	Stasiun BMKG	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (hari)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (hari)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (hari)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (hari)	Jumlah Curah Hujan (mm)	Jumlah Hari Hujan (hari)
Aceh	Sultan Iskandar Muda	1268.00	150.00	1098.00	137.00	1623.60	151.00	2264.40	142.00	1575.00	146.00
Sumatera Utara	Kualanamu ¹	2042.00	225.00	3175.00	227.00	2627.00	218.00	2148.00	200.00	975.90	105.00
Sumatera Barat	Sicinjin			4339.00	230.00	4627.40	232.00	2838.40	163.00	3548.00	185.00
Riau	Sultan Syarif Kasim II	2405.00	211.00	2636.00	217.00	2628.70	214.00	2343.70	188.00	2048.30	140.00
Jambi	Sultan Thaha ²	2295.00	209.00	1874.00	191.00	2093.60	229.00	1781.00	199.00	1694.90	135.00
Sumatera Selatan	Kenten	2593.00	217.00	3083.00	194.00	3409.20	238.00	1668.30	176.00	1947.20	138.00
Bengkulu	Pulau Baai			2545.00	143.00	3380.90	250.00	3323.00	174.00	2668.90	166.00
Lampung	Radin Inten II	1568.00	118.00	1685.00	143.00	2456.70	198.00	1682.50	178.00	1628.10	151.00
Kepulauan Bangka Belitung	Depati Amir	2921.00	213.00	2018.00	195.00	2839.20	243.00	1675.00	179.00	1534.70	163.00
Kepulauan Riau	Kijang	3893.00	205.00	3253.00	202.00	3389.40	227.00	3064.00	174.00	2250.90	174.00
DKI Jakarta	Kemayoran ³	1274.00		1570.00	139.00	2524.60	172.00	2908.00	157.00	2169.50	121.00
Jawa Barat	Bandung	1789.00	215.00	2510.00	219.00	2682.00	240.00	2388.00	226.00	2199.30	177.00

Dataset DKI Jakarta

Ini dataset yang diambil langsung dari website cuaca DKI Jakarta dari tahun 2011-2014.

Tabel 2 Cuaca DKI 2011

tahun	bulan	curah hujan(mm)	banyak hari hujan
2011	1	145,6	25
2011	2	230,7	18
2011	3	147,7	18
2011	4	106,8	16
2011	5	198,9	12
2011	6	70,5	10
2011	7	18,1	7
2011	8	1,5	2
2011	9	52,6	2
2011	10	80,1	12
2011	11	44,6	14
2011	12	177	17

Tabel 3 Cuaca DKI 2012

tahun	bulan	curah hujan (mm)	banyak hari hujan
2012	1	275,1	24,5
2012	2	157,9	16,8
2012	3	173,6	17,5
2012	4	196,2	14
2012	5	118	11
2012	6	67,2	4,5
2012	7	13,6	1,5
2012	8	2,4	1
2012	9	16,2	3,8
2012	10	44,3	5,3
2012	11	251,9	19
2012	12	254	20,3

Tabel 2 Cuaca DKI 2013

tahun	bulan	curah hujan (mm)	banyak hari hujan
2013	1	621,9	23
2013	2	146,6	20
2013	3	184,4	16
2013	4	204,3	18
2013	5	101	12
2013	6	256,7	19
2013	7	256,7	19
2013	8	61,4	8
2013	9	49,5	5
2013	10	110,1	9
2013	11	196,6	14
2013	12	338,9	23

Tabel 2 Cuaca DKI 2014

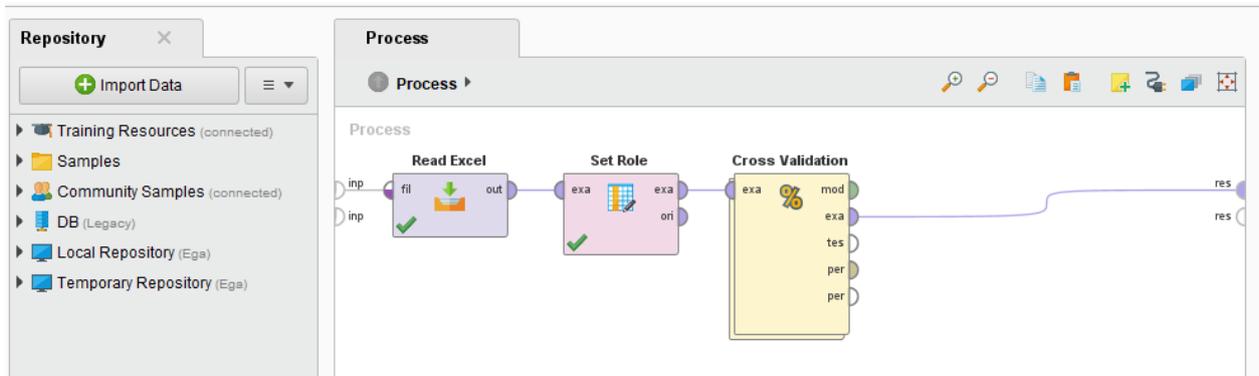
Tahun	Bulan	Curah Hujan (mm)	Banyak Hari Hujan
2014	1	1075	26
2014	2	689	22
2014	3	174	20
2014	4	168	16
2014	5	47	10
2014	6	174	12
2014	7	214	16
2014	8	39	4
2014	9	0	1
2014	10	52	4
2014	11	65	11
2014	12	211	15

3 Hasil dan Pembahasan

3.1 Hasil

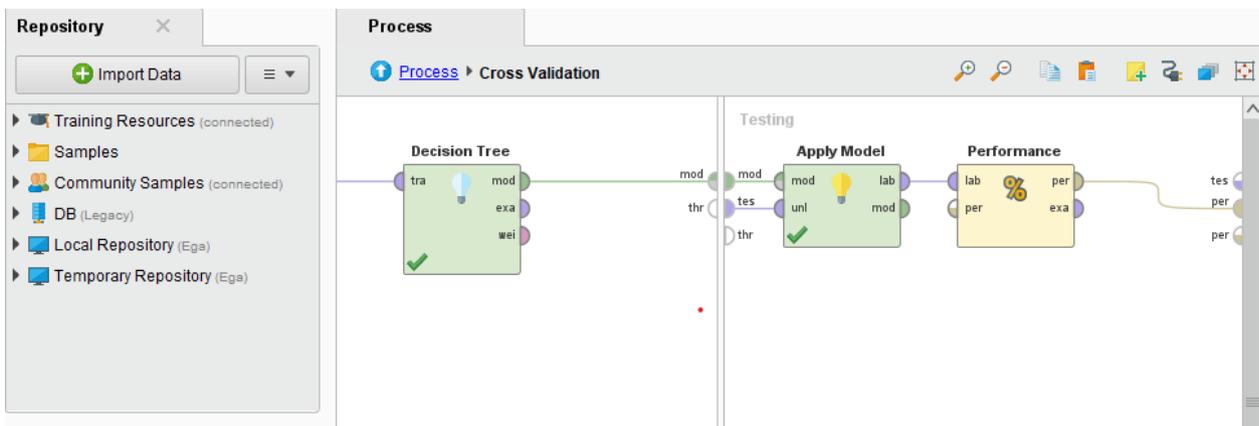
Dari data di atas kita bisa mendapatkan hasil untuk mensimulasikan dan mengklarifikasikan cuaca pada provinsi DKI JAKARTA untuk ke depannya. Menggunakan metode decision tree kita dapat melihat keefektivitasan dari data di atas sehingga kita bisa mengetahui apakah data tersebut bisa untuk mensimulasikan cuaca ke depannya.

Data yang digunakan dalam bentuk excel dibuka menggunakan read xls seperti gambar di bawah untuk memvalidasikan.



Selanjutnya setelah data dibuka melakukan tahap melatih dan menguji model dengan

menambahkan decision tree apply model dan performance. Seperti gambar di bawah



Selanjutnya kita mendapatkan hasilnya dari pengujian di atas yaitu 33,33% accuracy: 33.33%

	true Tahun	true Provinsi	true DKI Jakarta	class precision
pred. Tahun	1	1	1	33.33%
pred. Provinsi	0	0	0	0.00%
pred. DKI Jakarta	0	0	0	0.00%
class recall	100.00%	0.00%	0.00%	

3.2 Pembahasan

Pada data hasil di atas hanya mendapatkan 33,33% perisis nya dalam menggunakan metode decision tree. Sehingga masih ada cara yang lebih efektif untuk mensimulasikan dan mengklasifikasikan data yang sudah dapatkan. Jika kita gunakan untuk mensimulasikan mungkin kurang efektif menggunakan metode decision tree.

Dalam pembahasan disini suatu prakiraan dan pensimulasian dapat berjalan dengan metode decision tree. Tetapi keakuratannya masih belum begitu baik ketimbang metode lainnya yang ada.

4 Kesimpulan

Berdasarkan dari hasil di atas implementasi dan pengukuran modelnya dihasilkan data yang kurang presisi yaitu 33,33% sehingga bisa saja untuk mensimulasikannya atau

mungkin bisa menggunakan metode atau pemahaman yang lain. Bisa saja mensimulasikannya tetapi tidak akurat untuk klasifikasi digunakan.

5 Saran

Dalam data di atas sangat bagus tetapi kurang cocok untuk metode yang digunakan karena hanya mendapatkan tingkat presisi yang kecil untuk mensimulasikan prediksi cuacanya sangat kecil tingkat kevalidanya prediksi cuacanya. Jadi harus menggunakan suatu metode yang cocok agar dapat mensimulasikan yang akurat.

Referensi

- Dhawangkhar, M., & Riksakomara, E. (2017). Prediksi Intensitas Hujan Kota Surabaya dengan Matlab menggunakan Teknik Random Forest dan CART (Studi Kasus Kota Surabaya). *Jurnal Teknik ITS*, 6(1), 94-99.
- Gunadi, I. G., & Dewi, A. A. (2018). Klasifikasi Curah Hujan di Provinsi Bali Berdasarkan Metode Naïve Bayesian. *Wahana Matematika dan Sains: Jurnal Matematika, Sains, dan Pembelajarannya*, 12(1), 14-25.
- Laia, M. L., & Setyawan, Y. (2020). Perbandingan Hasil Klasifikasi Curah Hujan Menggunakan Metode SVM dan NBC. *Jurnal Statistika Industri dan Komputasi*, 5(2), 51-61.
- Lumbansiantar, D. S. (2019). Analisa Data Bencana Alam Untuk Prediksi Dampak Yang Ditimbulkan Dengan Algoritma J48 (Studi Kasus : Palang Merah Indonesia). *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1), 25-29. doi:10.30865/komik.v3i1.1562
- Maesaroh, S., & Kusri, K. (2017). Sistem Prediksi Produktifitas Pertanian Padi Menggunakan Data Mining. *Energy : Jurnal Ilmiah Ilmu-Ilmu Teknik*, 7(2), 25-39.
- Mendrofa, Y. (2019). Implementasi Algoritma C4.5 Untuk Memprediksi Tingkat Kerusakan Akibat Banjir (Studi Kasus : BPBD Prov.Sumut). *Jurnal Pelita Informatika*, 7(4), 584-592.
- Mujiasih, S. (2011). Pemanfaatan Data Mining Untuk Prakiraan Cuaca. *Jurnal Meteorologi Dan Geofisika*, 12(2), 189-195. doi:10.31172/jmg.v12i2.100
- Rosdiana, D., & Rismayana, A. H. (2018). Prediksi Waktu Tanam Cabai Menggunakan Algoritma C4.5. *Prosiding SINTAK* (pp. 436-442). Semarang: Universitas Stikubank.
- Saifudin, A. (2018). Metode Datamining untuk Seleksi Calon Mahasiswa Baru. *Jurnal Teknologi*, 10(1), 25-36. doi:10.24853/jurtek.10.1.25-36

- Siregar, A. M., Faisal, S., Cahyana, Y., & Priyatna, B. (2020). Perbandingan Algoritme Klasifikasi Untuk Prediksi Cuaca. *Accounting Information System*, 15-24.
- T., P. B., & Sari, R. D. (2017). Penerapan Data Mining untuk Prakiraan Cuaca di Kota Malang Menggunakan Algoritma Iterative Dichotomiser Tree (ID3). *JOUTICLA*, 3(2), 101-108.