

## Weather Analysis of Flight Delays at Husein Sastranegara Airport

Yedi Dermadi	Shinta Devi Lukitasari	Annisaa Nurhayati
Sekolah Teknik Elektro dan Informatika	Sekolah Teknik Elektro dan Informatika	Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung	Institut Teknologi Bandung	Institut Teknologi Bandung
Bandung, Indonesia	Bandung, Indonesia	Bandung, Indonesia
<a href="mailto:yedi.dermadi@students.itb.ac.id">yedi.dermadi@students.itb.ac.id</a>	<a href="mailto:shinta.lukitasari@students.itb.ac.id">shinta.lukitasari@students.itb.ac.id</a>	<a href="mailto:annisaa.nurhayati@students.itb.ac.id">annisaa.nurhayati@students.itb.ac.id</a>

*Abstract— Flight is an activity that is very vulnerable to weather conditions. The accuracy of weather information strongly supports flight activities. The effects of bad weather on flights include flight delays and flight cancellations. Based on data on flight delays from the Directorate General of Air Transportation of the Ministry of Transportation from January to March 2019 at Husein Sastranegara Airport, it is known that 20-30% of flight delays are caused by weather constraints. To estimate flight delays based on weather forecasts, weather data analysis is carried out to determine the type of weather that is endangering flights and causing flight delays. The analysis was carried out using the K-NN and Random Forest algorithms.*

*Keywords—weather, flight delay, kNN, random forest, tree, visualization*

### I. PENDAHULUAN

Penerbangan merupakan aktivitas yang sangat rentan terhadap kondisi cuaca. Peran cuaca dalam penerbangan sangat besar. Di satu sisi, kondisi cuaca bisa memberi dampak ekonomis berupa efisiensi pada pesawat terbang dan keselamatan penerbangan, di sisi lain cuaca memberi ancaman pada keselamatan penerbangan. Ketepatan informasi cuaca sangat mendukung kegiatan penerbangan.

Selain cuaca, iklim juga memiliki pengaruh terhadap operasional pesawat untuk penerbangan jauh. Saat perencanaan sebuah bandara aka dipelajari kondisi iklim pada lokasi bandara. Umumnya kondisi iklim yang dibutuhkan adalah faktor angin untuk mengetahui arah angin yang terbanyak dan untuk menghindari *cross wind* atau angin silang yaitu angin yang memotong arah landasan.

Dampak dari cuaca buruk terhadap penerbangan salah satunya adalah keterlambatan penerbangan dan pembatalan penerbangan. Keterlambatan penerbangan adalah terjadinya perbedaan waktu antara waktu keberangkatan atau kedatangan yang dijadwalkan dengan realisasi waktu keberangkatan atau kedatangan, sementara pembatalan penerbangan adalah tidak beroperasinya suatu penerbangan sesuai rencana penerbangan yang telah ditentukan [1]

Faktor yang menyebabkan keterlambatan penerbangan meliputi faktor cuaca, faktor manajemen airline, faktor teknis operasional, dan faktor lain-lain. Faktor cuaca meliputi hujan lebat, banjir, petir, badai, kabut, asap, jarak pandang di bawah standar minimal,

atau kecepatan angin yang melampaui standar maksimal yang mengganggu keselamatan penerbangan ditentukan [1].

Berdasarkan VFR (*Virtual Flight Rule*), pesawat tidak boleh lepas landas dan mendarat jika *ceiling* yang dilaporkan kurang dari 1.000 kaki dan jarak pandangnya tidak kurang dari tiga mil statute (4,8 km). *Ceiling* adalah ketinggian di atas permukaan bumi dari lapisan awan terendah. Untuk pesawat terbang dan amfibi, komponen lintas kecepatan angin yang terbukti aman untuk lepas landas dan mendarat adalah 90 derajat dan kecepatan angin harus setidaknya 20 knot atau 0,2 VS0, tidak melebihi 25 knot.

Berdasarkan data dari Kementerian Perhubungan Direktorat Jenderal Perhubungan Udara, diketahui jumlah keterlambatan penerbangan karena faktor cuaca di Bandara Husein Sastranegara yang dapat dilihat pada Tabel 1.

Tabel 1.

Data Keterlambatan Penerbangan Januari-Maret 2019 di Bandara Husein Sastranegara

No	Maskapai	Keterlambatan Berdasarkan Cuaca	Pembatalan Berdasarkan Cuaca
1	PT. Garuda Indonesia, TBK	5%	0%
2	PT. Lion Mentari Airlines	5%	0%
3	PT. Wings Abadi Airlines	8%	0%
4	PT. Indonesia Air Asia	1%	0%
5	PT. Citilink Indonesia	4%	0%
6	PT. Travel Express	1%	0%
7	PT. Nam Air	1%	0%

Oleh karena itu, untuk memperkirakan keterlambatan penerbangan berdasarkan prakiraan cuaca, dilakukan analisis data cuaca untuk mengetahui tipe cuaca yang membahayakan penerbangan dan mengakibatkan keterlambatan penerbangan. Analisis dilakukan menggunakan algoritma K-NN, Random Forest dan Tree.

## II. PENELITIAN TERKAIT

Terdapat penelitian-penelitian yang berfokus pada permasalahan keterlambatan penerbangan menggunakan data cuaca, diantaranya sebagai berikut.

Penelitian yang dilakukan oleh Choi dkk. [2] memprediksi keterlambatan maskapai yang disebabkan oleh kondisi cuaca buruk menggunakan data mining dan algoritma supervised machine learning. Data penerbangan domestik AS dan data cuaca dari 2005 hingga 2015 diekstraksi dan digunakan untuk model latih. Untuk mengatasi dampak data latih yang tidak seimbang, teknik pengambilan sampel diterapkan. Pohon keputusan, Random forest, AdaBoost dan k-Nearest-Neighbors diimplementasikan untuk membangun model yang dapat memprediksi keterlambatan penerbangan individu, kemudian masing-masing dari akurasi algoritma prediksi dan kurva *Receiver Operating Characteristic* (ROC) dibandingkan. Pada langkah prediksi, jadwal penerbangan dan prakiraan cuaca dikumpulkan dan dimasukkan ke dalam model. Dengan menggunakan data tersebut, model yang teruji melakukan klasifikasi biner untuk memperkirakan apakah penerbangan yang dijadwalkan akan terlambat atau tepat waktu.

Penelitian [3] menyajikan sebuah model estimasi *Departure Rate Decline* (DRD) berdasarkan analisis empiris disajikan untuk estimasi penurunan kapasitas bandara udara dalam berbagai kondisi cuaca. Kemudian, masalah pengambilan keputusan untuk pembatalan penerbangan keberangkatan dirumuskan sebagai model pemrograman linier integer dengan pertimbangan efisiensi bandara global dan prioritas maskapai penerbangan. Studi eksperimental menggunakan data nyata dari Beijing Capital International Airport menunjukkan kemanjuran metode yang diusulkan.

[4] melakukan penelitian untuk menganalisis keterlambatan penerbangan karena cuaca di Bandara Ahmad Yani Semarang dengan algoritma C4.5. Teknik ini menggunakan data cuaca dan daftar penerbangan penundaan penerbangan pada tahun 2013 di Bandara

Ahmad Yani di Semarang, sehingga data tersebut dapat digunakan sebagai referensi untuk memprediksi apakah akan ada penundaan penerbangan atau tidak. Rekomendasi kelayakan penerbangan dipengaruhi oleh arah angin, angin kecepatan, suhu, tekanan udara, jarak pandang, cuaca. Dan diperoleh hasil bahwa algoritma C4.5 diterapkan pada set data penerbangan Delay pada tahun 2013, data menghasilkan nilai akurasi matriks kebingungan 94,55% dan akurasi dalam interval 0,815 AOC 0 detik.

### III. METODE

Untuk memperkirakan tipe cuaca yang mempengaruhi keterlambatan atau pembatalan penerbangan, kami melakukan analisis terhadap data cuaca penerbangan di Bandara. Dari hasil analisis, akan diketahui cuaca pada waktu tertentu masuk ke dalam kategori ‘Aman’, ‘Sedang’, atau ‘Bahaya’. Pengkategorian dapat dilihat pada Tabel 2.

*Tabel 2*  
*Pengkategorian Cuaca.*

Kategori	Indikator	Keterangan
Aman	Jenis cuaca	
	Jarak pandang	
	Kecepatan Angin	
Sedang	Jenis cuaca	
	Jarak pandang	
	Kecepatan Angin	
Bahaya	Jenis cuaca	Hujan disertai guntur
	Jarak pandang	
	Kecepatan Angin	

Data yang telah diklasifikasikan akan disandingkan dengan data keterlambatan penerbangan sehingga dapat diketahui keterhubungan informasi cuaca dengan keterlambatan penerbangan.

#### A. Teknik Pengumpulan Data

##### 1) Data cuaca

Data cuaca didapatkan dari Kementerian Perhubungan Direktorat Jenderal Perhubungan Udara dan Pusat Meteorologi Penerbangan Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Data cuaca merupakan hasil pengumpulan data yang dilakukan BMKG dengan menggunakan sensor-sensor yang ditempatkan di lokasi tertentu untuk mengukur kondisi cuaca dalam berbagai parameter. Selain menghasilkan data pengukuran real time tentang kondisi cuaca saat itu, BMKG juga menghasilkan data prediksi kondisi cuaca pada beberapa waktu ke depan.

Pengamatan cuaca dilakukan setiap jam, namun untuk pelaporan dan pertukaran data cuaca secara internasional, dilakukan setiap 3 jam sekali, yaitu jam 00,03,06,09,12,15,18,21 UTC. Untuk wilayah WIB jam 00 UTC adalah jam 07.00 WIB. Pengamatan Cuaca dilakukan sejak H-10 jam pengamatan, unsur-unsur yang diamati meliputi Suhu Udara, titik embun, tekanan udara, jumlah curah hujan, keadaan cuaca, Awan, dan juga jarak pandang terjauh (*Visibility*). Dalam melakukan pengamatan cuaca, digunakan alat-alat meteorologi, seperti Sangkar Meteorologi, Barometer, Anemometer, Alat Penangkar Hujan (Ombrometer), Campbell Stokes, dan Panci Penguapan. Alat-alat tersebut diletakkan pada suatu tempat yang disebut Taman Alat. Setiap bandara pasti memiliki taman alat. Taman Alat berisi beberapa alat untuk mengamati unsur cuaca di tempat tersebut [5].

2) *Data keterlambatan penerbangan*

Data yang didapatkan merupakan data rekapitulasi laporan rutin masing-masing maskapai yang dihimpun oleh pihak Angkasa Pura dan dilaporkan kepada Direktorat Jenderal Perhubungan Udara Kementerian Perhubungan.

**B. Deskripsi Data**

Data yang didapatkan dari Kementerian Perhubungan Direktorat Jenderal Perhubungan Udara berupa data rekapitulasi waktu ketepatan terbang atau On Time Performance (OTP) dari setiap maskapai di Bandara Husein Sastranegara per bulan pada periode waktu Januari – Maret 2019. Deskripsi data dapat dilihat pada Tabel 3.

*Tabel 3.*

*Deskripsi Data dari Kementerian Perhubungan*

No.	Nama Kolom	Keterangan
1.	Maskapai	
2.	Total Penerbangan	
3.	Jumlah Tepat Waktu	
4.	Persentase Tepat Waktu	
5.	Jumlah Keterlambatan	
6.	Persentase Keterlambatan	
7.	Jumlah Pembatalan	
8.	Persentase Pembatalan	
9.	Jumlah Keterlambatan Berdasarkan Kategori Keterlambatan	I. (16-30) Menit TO NTO CUA LAIN II. (31-120) Menit TO NTO CUA LAIN III. (121-240) Menit TO NTO CUA LAIN IV. (>240) Menit TO NTO CUA LAIN
10.	Jumlah Pembatalan	TO NTO CUA LAIN

Data yang didapatkan dari Pusat Meteorologi Penerbangan berupa data cuaca penerbangan di Bandara Husein Sastranegara periode waktu Januari – Maret 2019. Deskripsi data dapat dilihat pada Tabel 4. Jumlah data sebanyak 994 record.

*Tabel 4.*

*Deskripsi Data dari Pusat Meteorologi Penerbangan*

No.	Nama Kolom
1.	Waktu pengisian
2.	Waktu pengamatan
3.	Angin
4.	Jarak pandang
5.	Cuaca

6.	Awan
7.	Suhu
8.	Titik embun
9.	Tekanan udara
10.	Jenis trend

### C. Pemrosesan Data

Tahap pertama yang dilakukan adalah mencari batas ambang dari faktor cuaca yang menentukan apakah penerbangan ditunda atau dibatalkan. Lalu, melakukan pelabelan data cuaca berdasarkan batas ambang tersebut. Pelabelan yang diberikan yaitu ‘Aman’, ‘Sedang’, dan ‘Bahaya’. Sedangkan data keterlambatan penerbangan dinormalisasi dengan mengabaikan informasi durasi keterlambatan dan menghapus data keterlambatan penerbangan yang bukan disebabkan oleh cuaca.

Setelah memiliki sejumlah data, data dibagi ke dalam 2 golongan. Yaitu data *training* sebesar 80% dan data *testing* sebesar 20%. Data *training* digunakan untuk membuat model klasifikasi. Sedangkan data *testing* digunakan untuk mengevaluasi hasil klasifikasi.

### D. Teknik Analisis Data

#### 1) Teknik Klasifikasi

Teknik klasifikasi yang kami gunakan adalah *supervised learning*. *Supervised learning* adalah sebuah pendekatan dimana sudah terdapat data yang dilatih, dan terdapat variable yang ditargetkan sehingga tujuan dari pendekatan ini adalah mengelompokkan suatu data ke data yang sudah ada.

#### 2) Algoritma

Algoritma yang kami gunakan adalah *K-Nearest Neighbor* (K-NN), *Random Forest*, dan *Tree*. Algoritma K-NN adalah sebuah metode klasifikasi terhadap sekumpulan data berdasarkan pembelajaran data yang sudah terklasifikasi sebelumnya. Termasuk dalam *supervised learning*, dimana hasil query instance yang baru diklasifikasikan berdasarkan mayoritas kedekatan jarak dari kategori yang ada dalam K-NN. Algoritma K-NN sangat sederhana, bekerja berdasarkan jarak terdekat antara data uji dengan data latih untuk mengetahui tetangga terdekat. KNN memiliki beberapa kelebihan, yaitu fleksibel dengan penambahan data dan efektif ketika data latih besar. Sementara kelemahan algoritman ini jika ada data uji yang sedang diuji, maka data uji akan tetap dikenali sebagai anggota kelas di database latih [6].

Algoritma *Random Forest* (RF) merupakan algoritma yang dapat meningkatkan hasil akurasi, karena dalam membangkitkan simpul anak untuk setiap *node* dilakukan secara acak. Metode ini digunakan untuk membangun pohon keputusan yang terdiri dari *root node*, *internal node*, dan *leaf node* dengan mengambil atribut dan data secara acak sesuai ketentuan yang diberlakukan. *Root node* merupakan simpul yang terletak paling atas, atau biasa disebut sebagai akar dari pohon keputusan. *Internal node* adalah simpul percabangan, dimana node ini mempunyai output minimal dua dan hanya ada satu input. Sedangkan *leaf node* atau *terminal node* merupakan simpul terakhir yang hanya memiliki satu input dan tidak mempunyai output. Pohon keputusan dimulai dengan cara menghitung nilai *entropy* sebagai penentu tingkat ketidakmurnian atribut dan nilai *information gain* [7].

#### 3) Evaluasi Rancangan

Evaluasi model yang dihasilkan dilakukan menggunakan data testing. Data testing merupakan 20% dari keseluruhan data. Parameter evaluasi diperoleh dari nilai akurasi model dalam mengklasifikasi data test tersebut.

## IV. HASIL DAN PEMBAHASAN

### A. Analisis Data

Data cuaca dibersihkan dari informasi yang tidak digunakan dan hanya menyisakan informasi yang dibutuhkan untuk menentukan kategori keamanan cuaca terhadap penerbangan. Variabel yang digunakan sebagai parameter berjumlah 7, yaitu:

- Arah angin
- Kecepatan angin
- Visibility
- Cuaca
- Suhu
- Titik embun
- Tekanan udara

Data keterlambatan penerbangan dibersihkan dari informasi keterlambatan yang tidak disebabkan oleh cuaca, dan dijumlahkan nilai kejadiannya per bulan. Data penerbangan ini tidak dapat dikorelasikan secara langsung dengan data cuaca yang ada, karena data yang diperoleh hanya berupa rekapitulasi tidak terdapat informasi waktu kejadian keterlambatan penerbangan.

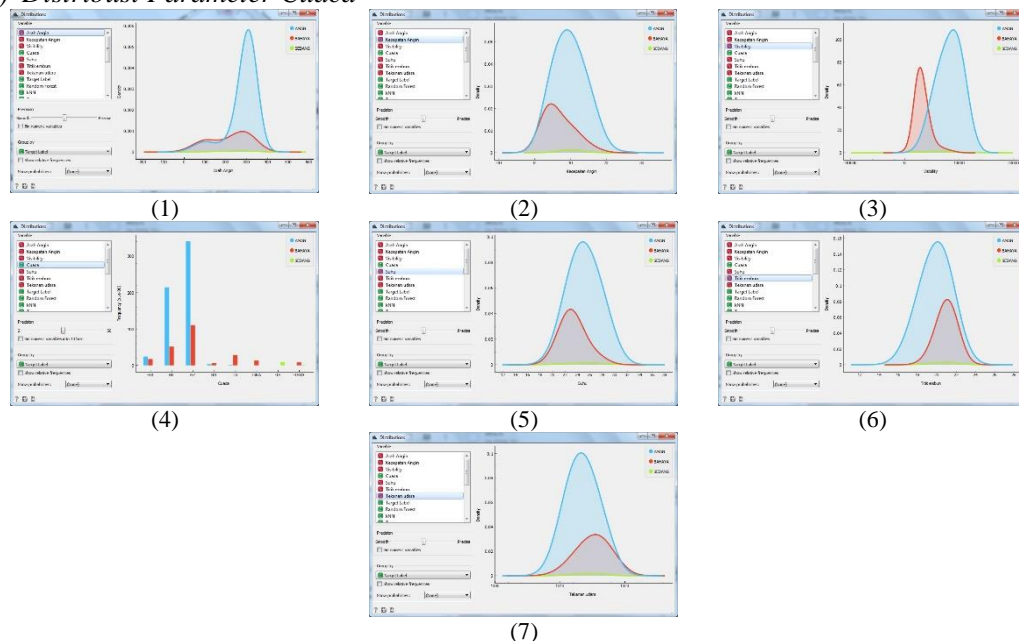
### B. Pembangunan Model yang Diusulkan

Model yang dibangun adalah model prediksi keamanan cuaca untuk penerbangan sesuai nilai parameter yang dimasukkan. Model dibangun dengan menggunakan algoritma *k-Nearest Neighbor*, *Random Forest*, dan *Decision Tree* dengan bantuan *Orange Analytic Tools* dengan menggunakan data latih yang telah disiapkan sebelumnya.

### C. Model Visualisasi

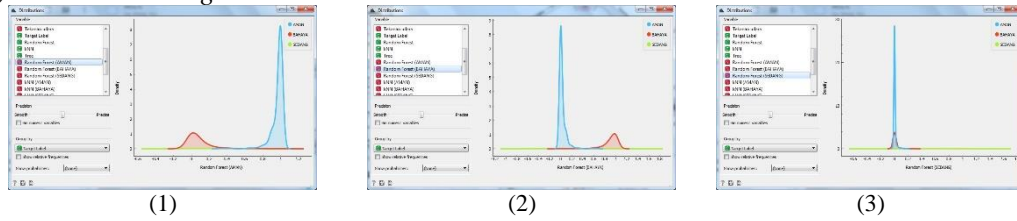
Visualisasi model yang dihasilkan *Orange Analytic Tools*, terdiri dari visualisasi distribusi parameter cuaca dan distribusi kategori ‘aman’, ‘sedang’, dan ‘bahaya’ menggunakan tiga algoritma yang dapat dilihat pada gambar-gambar dibawah ini.

#### 1) Distribusi Parameter Cuaca

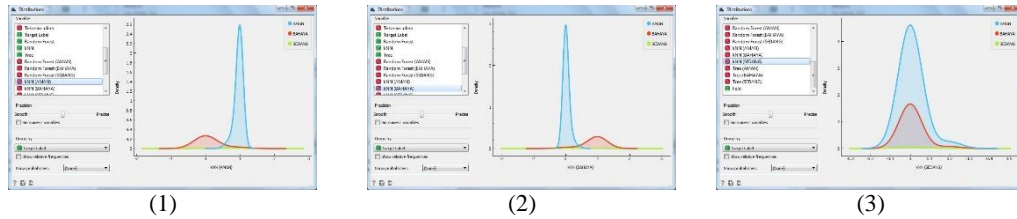


Gambar 1. Distribusi parameter cuaca. (1) Arah angin; (2) Kecepatan angin; (3) Visibility; (4) Tipe cuaca; (5) Suhu; (6) Titik embun; dan (7) Tekanan udara.

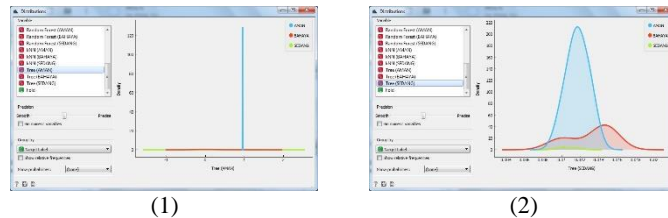
## 2) Distribusi Kategori Cuaca



Gambar 2. Distribusi kategori cuaca menggunakan Random Forest. (1) Kategori cuaca 'AMAN'; (2) Kategori cuaca 'BAHAYA'; dan (3) Kategori cuaca 'SEDANG'.

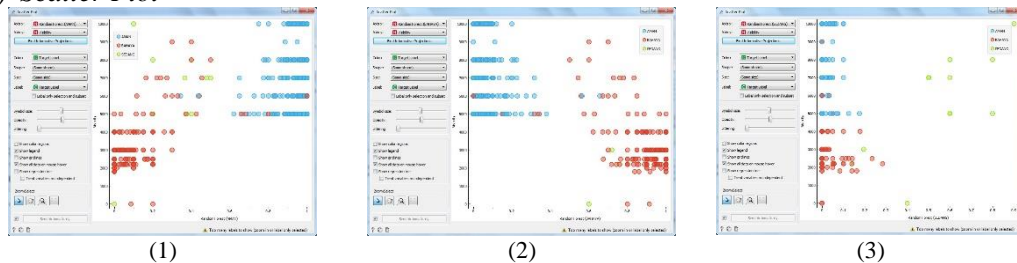


Gambar 3. Distribusi kategori cuaca menggunakan kNN. (1) Kategori cuaca 'AMAN'; (2) Kategori cuaca 'BAHAYA'; dan (3) Kategori cuaca 'SEDANG'..



Gambar 4. Distribusi kategori cuaca menggunakan Tree. (1) Kategori cuaca 'AMAN'; dan (2) Kategori cuaca 'SEDANG'.

## 3) Scatter Plot

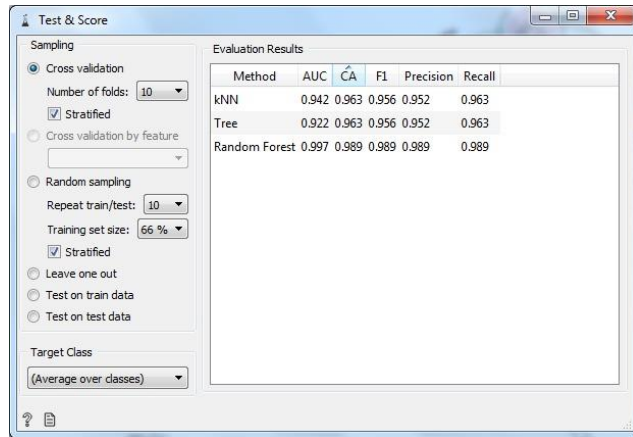


Gambar 5. Scatter Plot menggunakan Random Forest. (1) Kategori cuaca 'AMAN'; (2) Kategori cuaca 'BAHAYA'; dan (3) Kategori cuaca 'SEDANG';

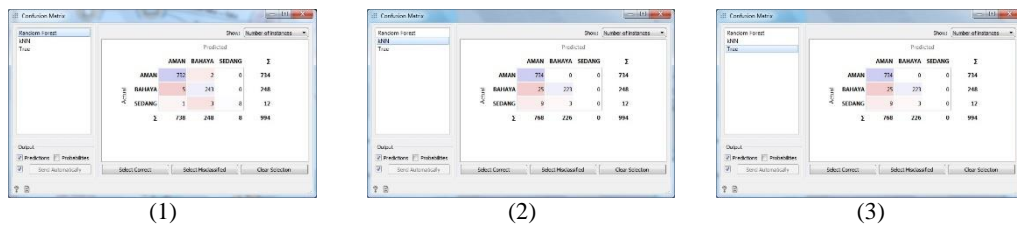
## D. Evaluasi Model

Model yang telah dihasilkan kemudian dievaluasi untuk mengetahui tingkat akurasi. Berdasarkan hasil evaluasi menggunakan Orange *Analytic Tools* pada Gambar 6. dapat diketahui bahwa akurasi model sebesar 0.963 dengan menggunakan kNN dan Tree dan 0.989 menggunakan Random Forest. Sementara akurasi hasil dapat dilihat pada confusion matrix yang dapat dilihat pada Gambar 7.





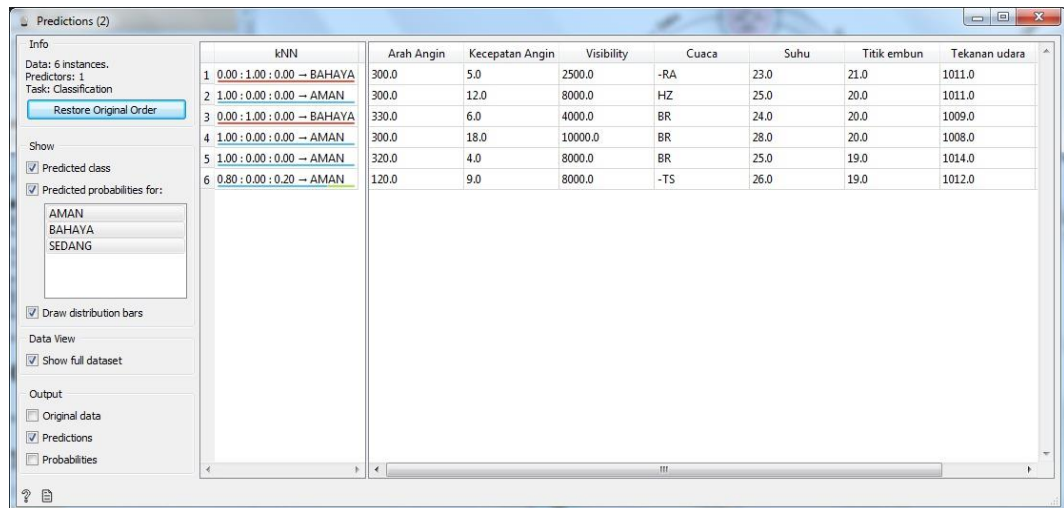
Gambar 6. Evaluasi Model



Gambar 7. Confusion matrix menggunakan (1) Random forest; (2) kNN; dan (4) Tree

## E. Analisis Prediksi

Untuk mengetahui apakah model yang dihasilkan telah sesuai, maka data uji yang telah disiapkan sebelumnya diinputkan ke dalam model. Hasilnya dapat dilihat pada Gambar 8, 9, dan 10.



Gambar 8. Hasil prediksi menggunakan kNN



	Random Forest	Arah Angin	Kecepatan Angin	Visibility	Cuaca	Suhu	Titik embun	Tekanan udara
1	0.05 : 0.95 : 0.00 → BAHAYA	300.0	5.0	2500.0	-RA	23.0	21.0	1011.0
2	1.00 : 0.00 : 0.00 → AMAN	300.0	12.0	8000.0	HZ	25.0	20.0	1011.0
3	0.30 : 0.70 : 0.00 → BAHAYA	330.0	6.0	4000.0	BR	24.0	20.0	1009.0
4	1.00 : 0.00 : 0.00 → AMAN	300.0	18.0	10000.0	BR	28.0	20.0	1008.0
5	1.00 : 0.00 : 0.00 → AMAN	320.0	4.0	8000.0	BR	25.0	19.0	1014.0
6	0.05 : 0.00 : 0.95 → SEDANG	120.0	9.0	8000.0	-TS	26.0	19.0	1012.0

Gambar 9. Hasil prediksi menggunakan Random Forest

	Tree	Arah Angin	Kecepatan Angin	Visibility	Cuaca	Suhu	Titik embun	Tekanan udara
1	0.00 : 0.99 : 0.01 → BAHAYA	300.0	5.0	2500.0	-RA	23.0	21.0	1011.0
2	0.96 : 0.03 : 0.01 → AMAN	300.0	12.0	8000.0	HZ	25.0	20.0	1011.0
3	0.00 : 0.99 : 0.01 → BAHAYA	330.0	6.0	4000.0	BR	24.0	20.0	1009.0
4	0.96 : 0.03 : 0.01 → AMAN	300.0	18.0	10000.0	BR	28.0	20.0	1008.0
5	0.96 : 0.03 : 0.01 → AMAN	320.0	4.0	8000.0	BR	25.0	19.0	1014.0
6	0.96 : 0.03 : 0.01 → AMAN	120.0	9.0	8000.0	-TS	26.0	19.0	1012.0

Gambar 10. Hasil prediksi menggunakan Tree

## V. KESIMPULAN

Berdasarkan data keterlambatan penerbangan dari Direktorat Jenderal perhubungan Udara Kementerian Perhubungan periode Januari s.d. Maret 2019 di Bandara Husein SASTRANEGARA diketahui bahwa 20 – 30% keterlambatan penerbangan disebabkan oleh kendala cuaca. Oleh karena itu penelitian ini dilakukan untuk menghasilkan sistem yang dapat memberikan prediksi keterlambatan penerbangan berdasarkan kondisi prakiraan cuaca. Dengan data cuaca yang dimiliki dan informasi mengenai kriteria keamanan cuaca untuk penerbangan dibuat model klasifikasi keamanan cuaca tersebut menggunakan algoritma *machine learning*. Berdasarkan model tersebut dapat dibuat prediksi keamanan cuaca untuk penerbangan dari parameter cuaca yang diberikan. Prediksi keterlambatan penerbangan belum dapat dilakukan dengan keterbatasan data yang diperoleh. Untuk dapat menghasilkan sistem yang lebih akurat, dibutuhkan data kejadian keterlambatan penerbangan yang lebih presisi untuk dapat diintegrasikan dengan data cuaca.

## REFERENSI

- [1] “Peraturan Menteri Perhubungan Republik Indonesia Nomor 9 Tahun 2015,” 2015.
- [2] S. Choi, Y. J. Kim, S. Briceno, dan D. Mavris, “Prediction of Weather-induced Airline Delays Based on Machine Learning Algorithms,” *2016 IEEE/AIAA 35th Digit. Avion. Syst. Conf.*, hal. 1–6, 2016.
- [3] X. Mao, Y. Yang, K. Q. Cai, dan W. H. Yang, “A decision support method for flight cancellations in adverse weather: An airport perspective,” *AIAA/IEEE Digit. Avion. Syst. Conf. - Proc.*, hal. 1E21-1E29, 2015.
- [4] M. N. Sholikhin dan Y. Rahayu, “Analisis Delay Penerbangan Akibat Cuaca di Bandara Ahmad Yani Semarang dengan Algoritma C4 . 5,” vol. 5, hal. 1–10, 2013.
- [5] Unknown, “Taman Alat Meteorologi.” [Daring]. Tersedia pada: <http://bagiankecilmeteorologi.blogspot.com/2014/01/taman-alat-meteorologi.html>.
- [6] A. D. Saputra, B. Irawan, dan R. A. Nugrahaeni, “DETECTION OF SKIN CANCER M ELANOM A USING EXPERT SYSTEM FORW ARD CHAINING M ETHOD AND IM AGE PROCESSING OF K -NEAREST NEIGHBOR ( KNN ) M ETHOD BASED ON ANDROID,” vol. 01, no. 02, hal. 27–36, 2018.
- [7] Y. S. Nugroho dan N. Emiliyawati, “Sistem Klasifikasi Variabel Tingkat Penerimaan Konsumen Terhadap Mobil Menggunakan Metode Random Forest,” *Sci. J. Informatics UNNES*, vol. 9, no. 1, 2017.