

## ANALISIS PREDIKSI BENCANA ANGIN PUYUH DI JAWA BARAT MENGUNAKAN ALGORITMA K-NN DAN C4.5

Yosep Nuryaman<sup>1</sup>, Afri Yudha<sup>2</sup> dan Ayuni Asistiyasari<sup>3</sup>

STMIK Cipta Karya Informatika<sup>1</sup>, Institute Muhammadiyah Bekasi<sup>2</sup>, Universitas BSI<sup>3</sup>

yosepnuryaman@gmail.com

### ABSTRAK

Salah satu efek dari perubahan iklim yang tidak disadari di negara Indonesia adalah adanya peningkatan terjadinya bencana angin puyuh. Berdasarkan data yang diambil dari BNPB Indonesia, kejadian bencana angin puyuh meningkat secara signifikan dalam 10 tahun terakhir. Tercatat pada tahun 2007 hanya 127 kejadian setahun, namun pada tahun 2017 terdapat 645 kejadian. Kerugian akibat kejadian bencana angin puyuh tidaklah kecil, bukan hanya kerugian material namun juga tercatat ratusan orang terluka bahkan belasan orang meninggal dunia. Oleh sebab itu kami mencoba mengklasifikasikan cuaca ketika terjadinya bencana angin puyuh menggunakan algoritma KNN dan C4.5 berbasis PSO di wilayah Jawa Barat. Tujuan penelitian adalah memberikan prediksi angin puyuh yang banyak terjadi di wilayah Jawa Barat. Dengan pendekatan atribut cuaca yang diambil dari data BMKG dan kejadian bencana angin puyuh yang diambil dari BNPB Indonesia. Penelitian ini mampu menghasilkan hasil terbaik dengan akurasi sebesar 83%, presisi 85%, recall 96,15% dan AUC 0,591 menggunakan algoritma KNN berbasis PSO. Sehingga algoritma KNN dapat digunakan sebagai alternatif rekomendasi bagi BNPB untuk penanggulangan bencana angin puyuh.

**Kata kunci:**C4.5, KNN, PSO, Angin Puyuh, Perubahan Iklim

### ABSTRACT

*One effect of climate change that is not realized in Indonesia is an increase in whirlwind disasters. Based on data taken from BNPB Indonesia, the incidence of whirlwinds has increased significantly in the last 10 years. In 2007 there were only 127 events a year, but in 2017 there were 645 incidents. Losses due to the whirlwind disaster were not small, not only material losses but also hundreds of people were injured and even dozens died. Therefore we try to classify the weather when a whirlwind disaster occurs using the KNN and C4.5 algorithm based on PSO in West Java. The purpose of this study is to provide predictions of quail angina which mostly occur in West Java. With the approach to weather attributes taken from BMKG data and whirlwind events taken from BNPB Indonesia. This study was able to produce the best results with an accuracy of 83%, 85% precision, 96.15% recall and AUC 0.591 using the PSN-based KNN algorithm. So that the KNN algorithm can be used as an alternative recommendation for BNPB for whirlwind disaster management.*

**Keyword:**C4.5, KNN, PSO, whirlwind disaster, climate change

### PENDAHULUAN

Sebagai negara kepulauan yang dihimpit 2 samudera yaitu samudera hindia dan samudera pasifik serta terletak di wilayah garis khatulistiwa, Indonesia memiliki kekayaan yang melimpah ruah baik kekayaan hewani maupun hayati. Namun dengan adanya perubahan iklim yang tidak menentu tentunya akan membawa efek negatif tersendiri bagi negara seperti Indonesia.

Dampak perubahan iklim yang semakin dirasakan di negara kepulauan seperti di Indonesia salah satunya yaitu adanya siklon tropis yang menghantui wilayah Indonesia. Bukan hanya itu saja, efek negatif lainnya yaitu tidak menentunya curah hujan dan meningkatnya bencana angin puyuh.



Sumber: BNPB (2018)

Gambar 1. Grafik Kejadian Angin Puyuh

Berdasarkan data yang diambil dari *website* Badan Penanggulangan Bencana Alam, kejadian angin puyuh meningkat cukup signifikan selama 10 tahun terakhir. Dari tahun 2007 yang hanya 127 kejadian selama 1 tahun, namun di tahun 2017 mencapai 645 kejadian. Kerugian yang diakibatkan bencana angin puyuh pun tidaklah sedikit, bukan hanya kerugian material namun juga tercatat ratusan orang luka-luka dan bahkan belasan orang meninggal dunia.

Tabel 1. Kerugian Angin Puyuh Tahun 2017

No	Kerugian	Jumlah
1	Menderita & Mengungsi	14.014 jiwa
2	Luka-luka	190 jiwa
3	Meninggal	16 jiwa
4	Rumah Rusak Ringan	7,175 unit
5	Rumah Rusak Sedang	2,112 unit
6	Rumah Rusak Berat	1,714 unit
7	Fasilitas Umum	130 unit

Sumber: BNPB (2018)

Pada prosesnya, peneliti belum menemukan penelitian tentang prediksi bencana angin puyuh di Indonesia. Penelitian sebelumnya diantaranya dilakukan oleh Novandya (2017), dalam penelitian menggunakan dataset yang dibentuk dari informasi yang dihasilkan situs peramalan cuaca yaitu World Weather Online terhitung sejak tanggal 12 Agustus 2016 pukul 01:00 sampai dengan 20 Agustus 2016 pukul 22:00. C4.5 mampu menghasilkan Akurasi dengan nilai sebesar 88.89%. namun hanya sebatas memprediksi cuacanya saja. Sedangkan pada penelitian yang dilakukan oleh Yunita (2017), algoritma Neural Network mampu menghasilkan akurasi 72,97% dalam memprediksi cuaca juga hanya terbatas pada prediksinya cuacanya saja.

Namun berdasarkan pembelajaran, peneliti mencoba melakukan pendekatan berdasarkan karakteristik terjadinya bencana angin puyuh di Indonesia menggunakan *Machine Learning*. Karakteristik yang dipelajari berupa keadaan curah hujan, kecepatan angin dan sebagainya.

## METODE

Menurut Larose dalam (Kusrini, dkk. 2009:8) Data mining adalah sebuah proses, yang mana dalam melakukan prosesnya harus sesuai dengan prosedur dari proses tersebut, yaitu CRISP-DM (Cross-Industry Standard Process for Data Mining), yang terdiri dari keseluruhan proses, preproses data, pembentukan model, model evaluasi, dan tahap akhir penyebaran model.



Sumber : Kusrini (2009)  
Gambar 2. Proses CRISP-DM

Adapun proses yang terjadi pada setiap fase dalam CRISP-DM dalam gambar 2 adalah sebagai berikut:

1. Fase Pemahaman Bisnis (*Business/Research Understanding Phase*). Kejadian Angin Puyuh memang pada dasarnya sulit diprediksi, Dengan pendekatan kriteria pada saat terjadinya angin puyuh, peneliti mencoba melakukan prediksi kriteria tersebut disaat yang akan datang, sehingga dimungkinkan kriteria-kriteria tersebut dapat memprediksi terjadinya bencana angin puyuh
2. Fase Pemahaman Data (*Data Understanding Phase*). Untuk pemilihan data diambil dari 2 sumber yaitu data keadaan cuaca saat terjadinya angin puyuh dari website BMKG sebanyak 3934 *record* dan data kejadian angin puyuh dari website BPBD.

Nama Stasiun	WMO ID	Tanggal	Suhu Minimum (°C)	Suhu Maksimum (°C)	Suhu Rata-rata (°C)	Kelembaban Rata-rata (%)	Curah Hujan (mm)	Lama Penyinaran (jam)	Kecepatan Angin Rata-rata (knot)	Arah Angin Terbanyak (deg)	Kecepatan Angin Terbesar (knot)	Arah Angin Saat Kecepatan Maksimum (deg)
Stasiun Klimatc	96753	01/01/2007	22	29.3	26	80	2	2.8	2	W	5	270
Stasiun Klimatc	96753	02/01/2007	22.2	30.7	25.9	77	2.7	3.4	2	W	5	315
Stasiun Klimatc	96753	03/01/2007	21	31.2	25.7	75	0	5.2	2	W	6	270
Stasiun Klimatc	96753	23/12/2017	22.2	31.4	26.1	83	8888	4.9	2	N	4	360
Stasiun Klimatc	96753	24/12/2017	23	32.6	27.3	76	0.3	3.2	9999	N	3	300
Stasiun Klimatc	96753	25/12/2017	20.1	9999	25.6	76	9999	10.3	1	N	3	360
Stasiun Klimatc	96753	26/12/2017	19.5	33	25.8	77	9999	10.8	2	N	8	270
Stasiun Klimatc	96753	27/12/2017	22.2	30.4	26	82	9999	9.5	1	N	4	270
Stasiun Klimatc	96753	28/12/2017	22	9999	9999	9999	9999	1.4	3	N	5	290
Stasiun Klimatc	96753	29/12/2017	20.5	32.2	9999	9999	9999	3.6	1	N	8	330
Stasiun Klimatc	96753	30/12/2017	23.5	31.9	26.2	84	9999	7.1	1	N	4	230
Stasiun Klimatc	96753	31/12/2017	22.6	32.4	25.9	84	3.3	3	1	N	4	270

Sumber : BMKG (2018)

Gambar 3. Keadaan Cuaca menurut BMKG

Dari data BMKG tersebut kemudian dikompilasikan dengan data BNPB untuk memberikan label kapan terjadinya bencana angin puyuh seperti pada gambar dibawah ini

Tanggal	Suhu Minimum (°C)	Suhu Maksimum (°C)	Suhu Rata-rata (°C)	Kelembaban Rata-rata (%)	Curah Hujan (mm)	Lama Penyinaran (jam)	Kecepatan Angin Rata-rata (knot)	Arah Angin Terbanyak (deg)	Kecepatan Angin Terbesar (knot)	Arah Angin Saat Kecepatan Maksimum (deg)	Kejadian Angin Puyuh
03/01/2007	21	31.2	25.7	75	0	5.2	2	W	6	270	ada
06/01/2007	21.2	32.6	26.9	70	0	5	1	N	9999	9999	ada
27/01/2007	23.5	30.8	25.8	93	7.5	2.1	1	N	9999	9999	ada
01/02/2007	23	27.3	24.6	93	20.6	0.5	9999	N	9999	9999	ada
02/02/2007	22.8	26	23.9	97	16.2	0	9999	W	1	270	ada
21/02/2007	21.7	29.8	25.3	84	1.6	2.3	9999	N	2	270	ada
08/04/2007	23.2	31.8	26.1	86	14.8	4.9	9999	S	3	180	ada
27/12/2017	22.2	30.4	26	82	9999	9.5	1	N	4	270	Tidak ada
28/12/2017	22	9999	9999	9999	9999	1.4	3	N	5	290	Tidak ada
29/12/2017	20.5	32.2	9999	9999	9999	3.6	1	N	8	330	Tidak ada
30/12/2017	23.5	31.9	26.2	84	9999	7.1	1	N	4	230	Tidak ada
31/12/2017	22.6	32.4	25.9	84	3.3	3	1	N	4	270	Tidak ada

Sumber : BNPB (2018)

Gambar 4. Penambahan Label Kejadian Angin Puyuh

3. Fase Pengolahan Data (*Data Preparation Phase*). Dari data yang telah diambil kemudian preproses dengan menghilangkan atribut yang tidak digunakan seperti data stasiun dan data tanggal. Hal kedua yang dilakukan adalah menghilangkan *record-record* yang tidak memiliki atribut lengkap. Dari hasil preprocessing maka data yang siap digunakan adalah sebanyak 160 data.

Tabel 2. Jumlah Data Siap Olah

No	Label	Jumlah Data
1	Ada	34
2	Tidak ada	126
Total		160

Sumber : Pengolahan data (2019)

4. Fase Pemodelan (*Modeling Phase*)  
Pemodelan yang dilakukan oleh peneliti yaitu menggunakan algoritma C4.5 dan KNN berbasis PSO untuk mencari pemodelan terbaik menggunakan *Rapid Miner*.
5. Fase Evaluasi (*Evaluation Phase*)  
Pada tahap ini dilakukan pengujian model untuk mendapatkan informasi model yang akurat. Evaluasi yang digunakan yaitu berupa *Confusion Matrix*.
6. Fase Penyebaran (*Deployment Phase*)  
Setelah pembentukan model dan evaluasi maka peneliti akan mengusulkan hasil penelitian yang dilakukan kepada pihak terkait dalam hal ini BPPD.

## HASIL

Setelah melakukan proses *preprocessing* data, peneliti menggunakan *Rapid Miner* untuk melihat hasil pemodelan serta akurasi, presisi dan *recall* setiap algoritma yang digunakan. Dan berikut hasil akurasi perhitungan Algoritma C4.5 berbasis PSO.

**accuracy: 72.50% +/- 9.76% (mikro: 72.50%)**

	true ada	true tidak ada	class precision
pred. ada	10	20	33.33%
pred. tidak ada	24	106	81.54%
class recall	29.41%	84.13%	

Sumber : Pengolahan Data (2019)

Gambar 3. Hasil Akurasi Algoritma KNN berbasis PSO

Dan berikut hasil perhitungan Algoritma KNN berbasis PSO

**accuracy: 83.12% +/- 4.88% (mikro: 83.12%)**

	true ada	true tidak ada	class precision
pred. ada	12	5	70.59%
pred. tidak ada	22	121	84.62%
class recall	35.29%	96.03%	

Sumber : Pengolahan Data (2019)

Gambar 4. Hasil Akurasi Algoritma KNN berbasis PSO

## SIMPULAN

Dari hasil perhitungan yang telah dilakukan dimana algoritma C4.5 berbasis PSO menghasilkan akurasi sebesar 72%, Presisi 82%, Recall 84% dibandingkan algoritma KNN berbasis PSO menghasilkan akurasi sebesar 83 %, Presisi 85%, Recall 96% dan Grafik AUC 0,591, maka dapat disimpulkan Algoritma KNN berbasis PSO merupakan algoritma terbaik yang mampu mengklasifikasikan kriteria-kriteria cuaca saat terjadinya dengan angin puyuh dengan akurasi sebesar 83 %, Presisi 85%, Recall 96% dan Grafik AUC 0,591.

## UCAPAN TERIMA KASIH

Ucapan terimakasih kepada BMKG dan BNPB yang telah memberikan aksesnya secara terbuka terhadap penelitian yang telah kami lakukan. Selain itu kami mengucapkan terimakasih kepada pihak STMIK CKI, Institute Muhamadiyah dan Universitas BSI yang telah mengizinkan dan mendukung kami dalam penulisan ini.

## DAFTAR RUJUKAN

- Abhishek. Kumar, Singh.M.P, Ghosh. S. dan Anand. A. (2012). Weather forecasting model using ANN. *Procedis Technology 4* (2012). 311-318
- Aprilia, Dennis, Donny Aji Baskoro, Lia Ambarwati dan I Wayan Simri Wicaksana. (2013). *Belajar Data Mining dengan Rapid Miner*. Jakarta: Gramedia Pustaka Utama.
- Badan Meteorologi, Klimatologi, dan Geofisika. (2018). Data Cuaca Harian 2007-2017. <http://dataonline.bmkg.go.id/home>. Diakses tanggal 24 Desember 2018.
- Badan Nasional Penanggulangan Bencana. (2018). Data Cuaca Harian 2007-2017. <http://dibi.bnpb.go.id/>. Diakses tanggal 24 Desember 2018.
- Kusrini dan Emha Taufiq Luthfi. (2009). *Algoritma Data Mining*. Yogyakarta: Andi Offset.
- Litta, A.J, Sumam Mary Idicula dan Mohanty. U.C. (2013). *Artificial Neural Network Model in Prediction of Meteorological Parameters during Premonsoon Thunderstroms*. Hindawi Publishing Corporation. 2013.
- Novandya, A., & Oktria, I. (2017). Penerapan Algoritma Klasifikasi Data Mining C4 . 5 Pada Dataset Cuaca Wilayah Bekasi, 6, 98–106.
- Yunita. (2015). Prediksi Cuaca Menggunakan Metode Neural Network. *Paradigma, XVII*(2), 47–53.