

BAB I PENDAHULUAN

1.1 Latar Belakang

Moda transportasi udara adalah moda transportasi jarak jauh yang sesuai untuk mengangkut barang maupun bahan mentah yang bernilai tinggi dan cepat rusak serta manusia dengan tingkat sensitivitas terhadap waktu yang tinggi (Tusmar & Mora, 2015). Menurut Pakan (2012), kelebihan-kelebihan yang dimiliki moda transportasi udara menjadikan moda transportasi ini sebagai moda transportasi dengan jumlah peminat lebih banyak dibandingkan dengan moda transportasi lain.

Seiring dengan perubahan waktu, penggunaan transportasi udara telah berkembang di seluruh dunia. Menurut laporan statistik dari *International Civil Aviation Organization* (ICAO), penggunaan transportasi udara terjadwal, baik penerbangan domestik maupun internasional di dunia, meningkat dari sebanyak 2,7 miliar penerbangan di tahun 2010 menjadi sebanyak 4,4 miliar penerbangan di tahun 2019 (ICAO, 2019). Pada tahun 2019, Indonesia berada pada urutan ke-23 negara dengan aktivitas penerbangan terbanyak di dunia (ICAO, 2019).

Munculnya pandemi *Coronavirus Disease* (COVID-19) sejak tahun 2019 sangat berdampak pada penurunan aktivitas penerbangan dunia. Khususnya di Indonesia, selama dua tahun terakhir banyak rute yang tidak dapat dioperasikan dikarenakan pembatasan mobilitas antar daerah maupun pembatasan Warga Negara Asing (WNA) masuk ke Indonesia (Kemenhub RI, 2021). Namun, seiring dengan pulihnya masa pandemi, jumlah penerbangan dan penumpang pesawat telah meningkat. Kementerian Perhubungan RI memprediksi terjadinya peningkatan jumlah penerbangan domestik di Indonesia pada tahun 2022, yaitu sebesar 9,97% dari tahun sebelumnya.

Rute penerbangan domestik di Indonesia terdiri dari banyak rute yang dapat diurutkan mulai dari rute tersibuk hingga rute yang tidak beroperasi dilihat dari jumlah penumpang. Menurut data yang dilansir dari laman OAG

(*Official Airline Guides*), salah satu rute tersibuk di Indonesia adalah rute penerbangan Jakarta menuju Denpasar. Pada tahun 2021, rute ini termasuk ke dalam 20 besar rute tersibuk di dunia dan pada tahun 2022 naik menjadi 10 besar menduduki peringkat ke-9 dengan banyak penumpang sebanyak 600.248 penumpang per September 2022 (OAG, 2020).

Dinamika di dalam industri transportasi udara menunjukkan bahwa seiring jumlah penumpang terus meningkat, tingkat keterlambatan penerbangan dapat berfluktuasi dari tahun ke tahun (Kim & Bae, 2020). Menurut Permenhub (2015), keterlambatan atau *delay* pada penerbangan terjadi karena adanya perbedaan antara waktu keberangkatan atau kedatangan yang dijadwalkan dengan waktu realisasinya. Kemenhub RI (2021) pun merangkum bahwa dari total 269.241 penerbangan oleh maskapai utama, 18,17% di antaranya mengalami *delay*. Selama tahun 2021, urutan lima maskapai utama dalam negeri dengan nilai *On Time Performance* (OTP) tertinggi 84,75% dan terendah 72,38% adalah PT. Garuda Indonesia, PT. Batik Air Indonesia, PT. Citilink, PT. Lion Mentari Airlines, dan PT. Indonesia Air Asia (Kemenhub RI, 2021). Nilai OTP mengindikasikan reputasi bagi suatu maskapai, di mana semakin tinggi nilai OTP, maka semakin baik reputasi maskapainya (Pakan, 2012).

Delay penerbangan dapat terjadi disebabkan oleh berbagai faktor. Faktor-faktor yang dapat menyebabkan terjadinya *delay* penerbangan terdiri dari Manajemen *Airlines*, Teknikal Operasional (TO), Non Teknis Operasional (NTO), Cuaca (CUA), dan faktor lainnya (Permenhub, 2015). Pada praktiknya, setiap penerbangan dapat berhadapan dengan kondisi tidak terduga yang mempengaruhi pelaksanaan jadwal seperti pengaruh cuaca. Menurut Permenhub No. PM 89 (2015), faktor cuaca dapat meliputi hujan lebat, badai, kabut, kelembaban, arah angin, kecepatan angin yang tinggi, suhu, dan kabut asap.

Delay penerbangan yang terjadi karena cuaca yang buruk dapat berpengaruh pada keamanan mesin pesawat. *Delay* penerbangan pun dapat menyebabkan banyak penumpang yang menumpuk di bandara bahkan tanpa

mendapat informasi yang memadai. Oleh karena itu, dalam rangka memenuhi kebutuhan informasi para penumpang pesawat, diperlukan analisis terhadap cuaca yang dapat memberikan informasi *delay* penerbangan.

Adanya peramalan cuaca sangat membantu dalam kegiatan penerbangan. Informasi cuaca pun dapat diakses oleh siapapun dari berbagai media informasi. Hal ini menjadikan banyak peneliti yang tertarik untuk meneliti *delay* penerbangan berdasarkan cuaca. Prasetyo (2019) menggunakan metode klasifikasi *decision tree* untuk menentukan *delay* penerbangan akibat cuaca pada berbagai maskapai. Penelitiannya menunjukkan tingkat nilai diagnosa klasifikasi yang baik dengan kurva *Receiver Operating Characteristic* (ROC) sebesar 0,996. Penelitian lainnya dilakukan oleh Dermadi dkk. (2019) dengan membandingkan metode klasifikasi algoritma K-NN, *Tree* dan *Random Forest* pada keterlambatan penerbangan dengan kendala cuaca. Hasil penelitian menunjukkan bahwa algoritma *Random Forest* memiliki akurasi model tertinggi yaitu sebesar 0,989.

Meskipun metode klasifikasi pada penelitian sebelumnya dapat menunjukkan hasil yang baik, penelitian tersebut belum ada yang mengikutsertakan variabel lamanya *delay*. Data lamanya *delay* dapat dianggap sebagai data *time-to-event*, yang berarti waktu hingga munculnya *event* berupa *delay*. Analisis yang dapat digunakan pada data *time-to-event* adalah analisis *survival* (Kleinbaum & Klein, 2005).

Penggunaan model di dalam analisis *survival* yang umum adalah model statistika multivariat *Cox Proportional Hazard* atau Cox PH (Deo dkk., 2021). Penelitian terdahulu oleh Kim dan Bae (2020) meneliti masalah *delay* keberangkatan pada penerbangan di Korea Selatan menggunakan analisis *survival* dengan model *Cox Proportional Hazard*. Hasil penelitian menunjukkan bahwa model tersebut dapat digunakan untuk mendeskripsikan *hazard ratio* serta mengidentifikasi faktor penyebab *delay* yaitu nama bandara dan slot waktu keberangkatan. Model Cox PH merupakan regresi Cox dengan aturan pemenuhan asumsi *proportional hazard* (Cox, 1972). Asumsi

proportional hazard menyebutkan bahwa fungsi *hazard* dari individu yang berbeda adalah proporsional atau rasio dari fungsi *hazard* dua individu yang berbeda adalah konstan (Lee & Wang, 2003).

Permasalahan muncul jika terdapat pelanggaran asumsi *proportional hazard* dalam data dimensi tinggi. Sebagai upaya untuk mengatasi masalah ini, metode non-parametrik dan fleksibel seperti pohon *survival* dan ensambel pohon telah dikembangkan, yaitu dengan metode *Random Survival Forest* (Saadati & Bagheri, 2019). *Random Survival Forest* (RSF) merupakan perluasan dari metode *Random Forest* (RF). RF diusulkan oleh Breiman pada tahun 2001 sebagai metode yang terfokus pada regresi dan klasifikasi (Ishwaran dkk., 2008). Metode RSF dikonstruksi dengan *bootstrap* pada pohon klasifikasi untuk data *survival* dan tidak memerlukan asumsi model (Saadati & Bagheri, 2019). Octapiani (2021) dalam penelitiannya membandingkan metode RSF dengan Cox PH pada data sistem berbagi sepeda, menunjukkan bahwa metode RSF lebih cocok digunakan untuk menganalisis data tersensor dengan melihat kriteria *C-index*.

Pohon klasifikasi dibentuk berdasarkan aturan pemisahan (*splitting rule*) untuk memecah *node* pohon menjadi *node* anak dengan probabilitas *survival* yang berbeda (Ishwaran dkk., 2008). Metode *Random Survival Forest* telah digunakan pada penelitian oleh Wanyonyi (2019) dengan data kematian balita di Kenya. Dalam penelitiannya, aturan pemisahan yang digunakan adalah *Log-Rank* dan *Log-Rank Score*. Hasil perbandingan antara dua aturan pemisahan tersebut menunjukkan bahwa *Log-Rank* memiliki tingkat prediksi yang lebih baik dengan *error rate* sebesar 5,35%.

Berdasarkan pemaparan di atas, peneliti tertarik untuk melakukan penelitian menggunakan penggabungan metode klasifikasi dan analisis *survival*, yaitu metode *Random Survival Forest* (RSF) dengan aturan pemisahan *Log-Rank* dan *Log-Rank Score*. Metode ini akan diterapkan pada data lima maskapai penerbangan yang mengalami *delay* penerbangan akibat cuaca pada rute bandara Soekarno Hatta (CGK) ke bandara I Gusti Ngurah Rai (DPS) terhitung mulai dari tanggal 1 Januari 2022 sampai dengan tanggal

5 Desember 2022. Adapun penelitian ini diberi judul “Metode *Random Survival Forest* dengan Aturan Pemisahan *Log-Rank* dan *Log-Rank Score* untuk Prediksi *Delay* Penerbangan Akibat Cuaca”. Hasil dari penelitian ini diharapkan dapat digunakan sebagai sumber yang fundamental untuk proses pengambilan keputusan di industri penerbangan.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan masalah yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. Bagaimana perbandingan tingkat kesalahan prediksi pada metode *Random Survival Forest* menggunakan aturan pemisahan *Log-Rank* dengan *Log-Rank Score*?
2. Bagaimana hasil prediksi *delay* penerbangan akibat cuaca menggunakan metode *Random Survival Forest* dengan aturan pemisahan terbaik?

1.3 Batasan Masalah

Dalam penulisan penelitian ini, penulis menentukan beberapa batasan masalah. Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data penelitian terfokus pada data *delay* penerbangan keberangkatan lima maskapai utama Indonesia, yaitu Lion Air, AirAsia, Batik Air, Garuda Indonesia dan Citilink pada rute bandara Soekarno Hatta menuju bandara I Gusti Ngurah Rai.
2. Pada metode *Random Survival Forest*, aturan pemisahan yang digunakan hanya *log-rank* dan *log-rank score*.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini berdasarkan rumusan masalah adalah sebagai berikut:

1. Menganalisis hasil perbandingan tingkat kesalahan prediksi pada metode *Random Survival Forest* menggunakan aturan pemisahan *Log-Rank* dengan *Log-Rank Score*.

2. Memaparkan hasil prediksi *delay* penerbangan akibat cuaca menggunakan metode *Random Survival Forest* dengan aturan pemisahan terbaik.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari hasil penelitian ini terbagi menjadi manfaat teoritis dan manfaat praktis yang diuraikan sebagai berikut:

1. Manfaat Teoritis

Hasil penelitian ini diharapkan dapat memberikan informasi mengenai analisis tingkat prediksi kesalahan untuk membandingkan aturan pemisahan *Log-Rank* dan *Log-Rank Score* pada metode *Random Survival Forest*, serta memberikan wawasan tentang dunia industri penerbangan.

2. Manfaat Praktis

Bagi pembaca khususnya perusahaan penerbangan di Indonesia, diharapkan hasil penelitian ini dapat menjadi referensi pengetahuan mengenai faktor cuaca yang menyebabkan terjadinya *delay* penerbangan sehingga masalah tersebut dapat ditangani dengan efektif.