

## BAB II

### TINJAUAN PUSTAKA DAN DASAR TEORI

#### 2.1 Tinjauan Pustaka

##### 2.1.1 Perkembangan Metode *Exponential Smoothing* dan *HWMS*

Beberapa studi telah dilakukan untuk menjelajahi berbagai metode dalam memprediksi lalu lintas pada jaringan telekomunikasi. Secara umum teknik prediksi *network traffic* terbagi menjadi 2 yaitu berbasis statistika dan *Machine Learning* (Lohrasbinasab dkk., 2022). Tran dkk. (2019) mengusulkan model sistem *forecasting* yang memiliki akurasi tinggi, komputasi yang tidak rumit, dan hemat biaya yaitu dengan metode *Exponential Smoothing*. Metode *Exponential Smoothing* yang ditawarkan adalah ditambahkannya aspek musiman pada model prediksi *time series* sebelumnya. Metode *Single*, *double*, *Holt-Winter's no seasonal*, *Holt-Winter's additive seasonal* dan *Holt-Winter's multiplicative seasonal (HWMS)* diujikan hingga menemukan hasil bahwa metode yang paling cocok untuk data set *hourly* adalah *HWMS* sedangkan untuk data set *daily*, baik *HWMS* ataupun *ETS* hasilnya terbilang tidak jauh. Tran menyebutkan bahwa komponen kunci dari perencanaan, pengembangan dan manajemen jaringan telekomunikasi adalah dengan memprediksi *traffic* nirkabel di masa mendatang. Sebuah sistem prediksi yang efisien harus memiliki akurasi tinggi, tingkat kompleksitas komputasi yang rendah dan tentu saja hemat biaya. Terbukti *exponential smoothing* merupakan metode yang cukup efektif digunakan untuk memprediksi *traffic*, namun keakuratan dan kesesuaian data *traffic* yang beragam masih dievaluasi (Tran dkk., 2019).

*Forecasting* menggunakan metode *exponential smoothing* memiliki tingkat akurasi prediksi yang lebih baik jika dibandingkan dengan metode lain. Makridakis dan Spiliotis (2018) melakukan benchmarking antara metode statistik tradisional antara lain *ARIMA*, *SES*, *Holt's*, dan *Exponential Triple Smoothing* dibandingkan dengan metode *machine learning* yang sedang menjadi tren seperti *LSTM*. Dataset yang digunakan adalah dataset yang sama persis digunakan pada kompetisi M3. Hasil dari *benchmarking* menunjukkan bahwa rata rata metode

statistik tradisional memiliki akurasi prediksi lebih tinggi dibandingkan algoritma *machine learning* yang sedang tren ( Makridakis dan Spiliotis, 2018). Bisa diartikan bahwa tidak selamanya metode tradisional kalah dengan *machine learning* modern dan masih layak untuk digunakan.

Kombinasi metode *Holt's Trend Exponential Smoothing* dan *K-mean* digunakan oleh Aldhyani dan Joshi untuk melakukan prediksi *traffic* guna menentukan kebutuhan bandwidth dimasa mendatang. Karena banyaknya *cell* dalam satu wilayah, maka akan sulit untuk melakukan prediksi satu persatu *cell* yang ada. Untuk menjawab permasalahan tersebut menggabungkan dua metode yaitu *clustering* dan statistika adalah solusinya. Metode *Clustering* yang digunakan adalah *K-means* sedangkan metode peramalannya menggunakan *Holt's Trend Exponential Smoothing*. Hasilnya model baru yang diusung memiliki akurasi peramalan yang lebih baik jika dibandingkan dengan metode *Holt's Trend Exponential Smoothing* tanpa adanya clustering dengan perbandingan *MAPE* dan *RMSE* model lama sebesar 0.0489 dan 2.8579 sedangkan model yang diusulkan 0.0319 dan 1.8107 (Aldhyani dan Joshi, 2017). Menggabungkan dua metode menjadi sebuah *framework* yang baik dan mampu menjawab permasalahan prediksi *traffic* sesuai kebutuhan.

Gabungan *Exponential Smoothing* dan *Fuzzy C-Means* oleh Aldhyani dkk., tahun 2020 digunakan sebagai metode smoothing pada tahap *preprocessing* dimana selanjutnya prediksi dilakukan secara *Deep Learning* menggunakan dua metode yaitu *Long-Short Time Memory (LSTM)* dan *Adaptive Neuro-Fuzzy Inference System (ANFIS)*. *Forecasting* dilakukan secara *sequel mining* untuk mendapatkan akurasi prediksi yang sangat tinggi. Hasilnya dengan menggabungkan metode *clustering*, *time series*, dan *Deep Learning* didapatkan hasil prediksi dengan tingkat akurasi yang lebih tinggi dibandingkan hanya menggunakan metode konvensional saja. Peningkatan akurasi prediksi dirasa penting karena saat ini pemanfaatan prediksi *traffic* untuk mendapatkan standar *QOS* yang diinginkan perusahaan telekomunikasi semakin tinggi. Dalam manajemen jaringan seluler, prediksi *traffic* tidak hanya berperan untuk mencegah *congestion*, akan tetapi juga bisa memaksimalkan sumberdaya yang ada dan

meningkatkan keamanan sebuah jaringan (Aldhyani dkk., 2020). Kelemahan penggabungan metode ini adalah memiliki komputasi yang rumit, namun memiliki kelebihan akurasi yang sangat tinggi.

### 2.1.2 Perkembangan Metode ARIMA

Metode tradisional dengan komputasi yang ringan sebenarnya masih layak digunakan memprediksi lalu lintas seluler seperti yang dilakukan (Syam dan Girsang, 2020). Ia menggunakan metode *Auto Regressive Integrated Moving Average (ARIMA)* dengan parameter (0.0.6) dalam memprediksi lalu lintas seluler dalam membantu dalam penyediaan bandwidth *LTE* dan hasilnya memuaskan. Apalagi bila metode tradisional ini diimprovisasi dengan menggabungkan dengan metode tambahan lain seperti yang dikerjakan oleh (Arifin dan Habibie, 2020). Arifin dan Habibi (2020) dalam penelitiannya menggabungkan *ARIMA* sebagai teknik statistik dikombinasikan dengan *disruptive equations* sebagai teknik *judgmental*-nya untuk mendapatkan hasil prediksi yang lebih akurat dibandingkan hanya menggunakan metode statistik saja untuk memprediksi *traffic LTE* (Arifin dan Habibie, 2020).

Penggabungan metode tersebut menjawab permasalahan pertumbuhan *traffic* tidak hanya disebabkan oleh evolusi jaringan seluler seperti *Long Term Evolution (LTE)* termasuk juga *Internet of Things* namun juga dipengaruhi oleh faktor Politik, Ekonomi, Sosial dan Teknologi (PEST). Kombinasi ini juga memberikan warna lain dalam memprediksi *traffic* yang lebih relevan yaitu selain hanya menggunakan teknik prediksi statistik juga diperlukan pendekatan *judgmental*. Akan tetapi *ARIMA* sangat sensitif dengan parameter urutan *autoregressive (p)*, urutan *differencing (d)*, dan urutan *moving average (q)*. Penggabungan *ARIMA* dengan *disruptive* formula memunculkan parameter tambahan yang perlu ditentukan dengan tepat. Menentukan parameter *disruptive* formula yang optimal dapat menjadi sulit jika ada perubahan baru atau pola yang tidak terduga, dan kesalahan dalam pemilihan parameter dapat mempengaruhi keakuratan prediksi.

### 2.1.3 Perkembangan Metode *FbProphet*

Metode *machine learning FbProphet* oleh Facebook saat ini juga sedang populer karena membantu operator seluler dalam melakukan perencanaan yang efisien dan menghemat biaya investasi karena pengaplikasiannya yang mudah jika dibandingkan *LSTM* (Chmieliauskas dan Gursnys, 2019). *FbProphet* telah digunakan Chmieliauskas dan Gursnys, (2019), Xu dkk., (2018), Chakraborty dkk., (2020), Rankothge dkk., (2022), Shuvo dkk., (2021), Owusu-kumih dkk., (2022), serta Cembaluk dan Aniszewski, (2022) untuk memprediksi pertumbuhan lalu lintas seluler pada jaringan telekomunikasi pada penelitian mereka. Tidak hanya *traffic*, Owusu-kumih dkk., (2022) juga memprediksi pendapatan yang dihasilkan oleh penyedia layanan telekomunikasi di Ghana menggunakan *FbProphet*.

*FbProphet* pernah dibandingkan akurasi dengan *ARIMA*, *ETS*, dan *SNAIVE* oleh Shuvo dkk., (2021). Hasilnya *FbProphet* berhasil mencatatkan akurasi 90% pada datasetnya. Selain itu, pada dataset *traffic* 5G Chakraborty dkk., (2020) memiliki kelebihan dalam bekerja pada data yang tidak lengkap, data outlier, perubahan mendadak dalam seri data, dan memberikan perkiraan yang masuk akal.

Xu dkk., (2018) pernah mengusung *Hybrid* model dari *FbProphet* dan *K-means*, konsepnya hampir sama dengan Aldhyani dan Joshi, (2017). *FbProphet* menjadi metode untuk memprediksi volume *traffic*, sedangkan *K-Means* sebagai algoritma untuk menentukan cluster dataset *holiday*. Dari analisa yang telah dilakukan membuktikan bahwa model *hybrid FbProphet* dan *K-Means* memiliki akurasi yang lebih baik saat melakukan peramalan lonjakan hari libur jika dibandingkan hanya menggunakan *FbProphet* saja (Xu dkk., 2018). Secara umum *FbProphet* mudah digunakan dan memiliki akurasi yang baik. Namun menggabungkan *FbProphet* dengan algoritma clustering tidak selalu relevan untuk semua dataset. Pendekatan ini harus didasarkan pada karakteristik dan pola data yang ada. Terkadang, model *FbProphet* dapat memberikan prediksi yang memadai tanpa memerlukan penggabungan dengan *K-means* seperti pada penelitian Rankothge dkk., (2022), Cembaluk dan Aniszewski, (2022), dan

Owusu-kumih dkk., (2022), sehingga *FbProhet* dipilih sebagai salah satu metode pembandingan pada penelitian kali ini.

## 2.2 Dasar Teori

### 2.2.1 *Time Series Analysis (Forecasting)*

*Forecasting* pada dasarnya adalah metode peramalan untuk mengetahui apa yang terjadi di masa depan dengan teknik analisa data tertentu sehingga sangat erat kaitannya dengan analisa data *time series*. Analisa *Time Series* merupakan salah satu metode yang mempelajari rangkaian deret waktu tertentu dimana dari apa yang terjadi pada deret waktu tersebut dijadikan pembelajaran untuk memprediksi nilai pada waktu yang akan datang. Hyndman dan Athanasopoulos, (2018) memaparkan *forecasting* merupakan pekerjaan untuk memprediksi masa depan seakurat mungkin dengan mempertimbangkan semua informasi yang ada. (Hyndman dan Athanasopoulos, 2018). *Forecasting* sangat erat kaitannya dengan pengambilan keputusan dan kebijakan dalam organisasi. Hasil dari *forecasting* menjadi peranan penting dimana akan digunakan sebagai bahan acuan untuk menentukan keputusan, perencanaan ataupun proyeksi perencanaan di masa depan. Selain bidang bisnis dan industri, *forecasting* juga sangat berguna untuk kebutuhan berbagai macam bidang organisasi seperti pemerintahan, sektor perekonomian, ilmu pengetahuan, kesehatan dan ilmu politik (Montgomery dan Jennings, 2007).

Mengetahui prediksi kejadian di masa depan akan menjadi masukan penting ke dalam banyak proses perencanaan dan pengambilan keputusan dalam area – area organisasi seperti berikut:

1. Manajemen Operasional. Organisasi sering menggunakan prediksi penjualan atau permintaan layanan untuk membuat jadwal produksi, penentuan syarat kepegawaian, kontrol inventaris dan kelola rantai pasokan. Hasil prediksi juga dapat digunakan untuk menentukan produk atau kombinasi layanan yang akan diberikan dan lokasi dimana produk tersebut akan diproduksi (Montgomery dan Jennings, 2007).
2. Marketing. *Forecasting* digunakan dalam banyak sekali penentuan keputusan dalam dunia pemasaran. Perubahan kebijakan, perkiraan respon dari iklan,

promosi, dan naik turunnya harga memungkinkan organisasi bisnis bisa mengevaluasi efektivitas dan arah sasaran sudah tepat atau perlu penyesuaian (Montgomery dan Jennings, 2007).

3. Keuangan dan Manajemen Resiko. Investor dalam aset keuangan membutuhkan prediksi kapan dan berapa keuntungan nilai investasi yang diberikan. Aset keuangan ini termasuk tetapi tidak terbatas pada saham, obligasi, dan komoditas saja. Keputusan investasi lainnya dapat dibuat dengan penyesuaian terhadap perkiraan suku bunga, opsi, dan nilai tukar mata uang. Manajemen risiko keuangan memerlukan estimasi volatilitas pengembalian aset sehingga risiko yang terkait dengan portofolio investasi dapat dinilai dan dijamin, dan derivatif keuangan dapat dinilai (Montgomery dan Jennings, 2007).
4. Ekonomi. Pemerintah, lembaga keuangan, dan organisasi pembuat kebijakan memerlukan prakiraan variabel ekonomi utama, seperti PDB, inflasi, suku bunga, pengangguran, pertumbuhan populasi, pertumbuhan lapangan kerja, produksi dan konsumsi. Prakiraan ini memainkan peran penting dalam pembahasan kebijakan moneter dan fiskal serta rencana anggaran dan keputusan pemerintah perencanaan strategis dan pengambilan keputusan organisasi bisnis dan lembaga keuangan (Montgomery dan Jennings, 2007).
5. Pengendalian Proses Industri. Memperkirakan nilai masa depan dari karakteristik kualitas kunci dari proses produksi dapat membantu menentukan kapan variabel penting yang dapat dikontrol dalam proses perlu diubah, atau apakah proses perlu ditutup dan dimodifikasi (Montgomery dan Jennings, 2007).
6. Demografi. Prakiraan populasi rutin dilakukan dan biasanya dibuat kelompok-kelompok berdasarkan wilayah, jenis kelamin, usia, ras, jumlah kematian dan kelahiran. Pemerintah menggunakan prediksi ini untuk menentukan kebijakan dan perencanaan yang bersifat sosial, sedangkan organisasi bisnis menggunakannya untuk rencana strategis terkait target pasar, pengembangan produk dan atau layanan yang ditawarkan (Montgomery dan Jennings, 2007).

Analisa *Time Series* dipengaruhi beberapa faktor yang berdampak pada akurasi prediksi antara lain data historis yang memiliki variabel *seasonality*, *trend*, dan *residual* juga pengetahuan saat ini dan pengetahuan tentang peristiwa di masa mendatang yang kemungkinan akan mempengaruhi prediksi. *Seasonality* merupakan pengaruhnya pergerakan data dikarenakan adanya faktor musiman seperti hari raya, musim libur sekolah dan sebagainya. Sedangkan untuk *trend* biasanya dipengaruhi oleh perkembangan terhadap suatu obyek yang berkaitan seperti *traffic* LTE setiap harinya semakin naik dikarenakan perkembangan teknologi *mobile phone*, sosial media dan bergesernya transaksi manual ke digital. Residual yang dimaksud mempengaruhi analisa *time series* merupakan data sisa yang muncul ketika sudah melakukan *fitting* model dalam analisis, karena harus diketahui residu tersebut berkorelasi dengan data atau tidak (Hyndman dan Athanasopoulos, 2018), (Montgomery dan Jennings, 2007).

### 2.2.2 *Holt-Winter's Multiplicative Seasonal (HWMS)*

*Holt-Winter's Multiplicative Seasonal* merupakan bagian dari keluarga metode prediksi *exponential smoothing*, dimana *exponential smoothing* itu sendiri adalah suatu metode prediksi data *time series* yang menimbang rata-rata observasi dari data-data masa lalu secara eksponensial, dimana pembobotan bergerak menurun seiring dengan semakin bertambahnya jarak observasi data yang ada di masa lalu. Atau bisa dikatakan semakin baru data observasi maka semakin tinggi pula bobotnya. Teknik ini dilakukan karena data terbaru adalah data yang seharusnya paling relevan memiliki bobot yang lebih besar. Parameter penghalusan (*smoothing*) biasanya dilambangkan dengan  $\alpha$  (*alpha*),  $\beta^*$  (*beta*),  $\gamma$  (*gamma*). Sedangkan *Holt-Winter's* adalah pemodelan dari *exponential smoothing* sederhana yang dikembangkan. *Holt-Winter's* menggunakan 3 konstanta pemulusan yaitu konstanta untuk pemulusan keseluruhan (*level*)  $l_t$ , pemulusan kecenderungan (*trend*)  $hb_t$ , dan pemulusan musiman (*seasonal*)  $S(t)$ .

Ada 2 pendekatan yang bisa dilakukan dengan *Holt-Winter's*, yaitu multiplikatif dan adaptif. Multiplikatif digunakan untuk data *time series* dengan variasi musiman yang sangat fluktuatif sedangkan adaptif lebih konstan

(Hyndman dan Athanasopoulos, 2018). Namun yang ditekankan dalam penelitian nantinya adalah pendekatan multiplikatif sesuai dengan persamaan 2.1.

$$y_{t+h|t} = (l_t + hb_t) * S(t + h - m(k + 1)) \quad (2.1)$$

dengan:

$y_{t+h|t}$  : Nilai prediksi di masa depan  $h$  periode setelah waktu ( $t$ ), yang dinyatakan sebagai ( $y$ ). Ini adalah apa yang akan di prediksi.

$l_t$  : Level pemulusan keseluruhan atau tingkat dasar pada waktu ( $t$ ).

$hb_t$  : Komponen kecepatan atau laju pertumbuhan pada waktu ( $t$ ).

$S$  : Nilai aktual komponen musiman pada waktu tertentu atau perkiraan nilai komponen musiman.

$t$  : Titik waktu atau periode yang sedang diamati atau yang ingin diprediksi. Ini adalah waktu saat ini atau titik referensi dari mana peneliti ingin melihat periode prediksi ke depan.

$S(t + h)$ : Komponen musiman pada waktu ( $t + h$ )

$h$  : Jumlah periode ke depan yang ingin diprediksi.

$m$  : Jumlah periode dalam satu siklus musiman.

$k$  : Jumlah siklus musiman yang telah terjadi hingga waktu  $t$

$k + 1$  : Langkah berikutnya dalam jumlah siklus musiman.

$S(t + h - m(k + 1))$  : Komponen ini digunakan untuk menangani pola musiman dalam data. Ini adalah sesuatu yang naik-turun dengan pola yang sama setiap ( $k$ ) periode. Ini adalah komponen musiman yang muncul dalam model peramalan.

Dalam metode HWMS, komponen seasonality bersifat relatif dan deret waktu dapat disesuaikan secara musiman dengan membaginya dengan komponen musiman. Level pemulusan ( $l_t$ ) Ini menggambarkan tingkat dasar atau rata-rata data pada titik waktu ( $t$ ). Ini adalah komponen dasar dalam model peramalan. Pemulusan kecenderungan (*trend*) ( $hb_t$ ) Ini menggambarkan sejauh mana data mengalami pertumbuhan atau penurunan. Ini memungkinkan model untuk mengakomodasi kecenderungan. Ekspresi  $S(t + h - m(k + 1))$  secara lengkap mengacu pada komponen musiman pada waktu ( $t + h$ ) dengan



mempertimbangkan pola musiman  $m$  langkah siklus musiman ( $k$ ) dan jumlah periode berikutnya.

### 2.2.3 *Rolling Forecast*

Dalam praktek *forecasting* ada istilah *training set* dan *test set*. *Training set* digunakan model untuk mempelajari kemudian menemukan pola data yang ada digunakan. Sedangkan *test set* berfungsi sebagai validasi model yang akan dipakai. Validasi penting dilakukan untuk mengetahui seberapa dekat prediksi yang dihasilkan oleh metode aslinya dalam hal ini *HWMS*. Besaran residual pada saat melakukan *training* bukan yang menjadi acuan model tersebut baik atau tidak, namun keakuratan ditentukan dengan mempertimbangkan seberapa baik model dapat menyesuaikan dengan data baru.

Melakukan *fitting* model umumnya membagi data yang ada menjadi 2 bagian, 80% sebagai *training* dan 20% sebagai *test* seperti pada ilustrasi Gambar 2.1.

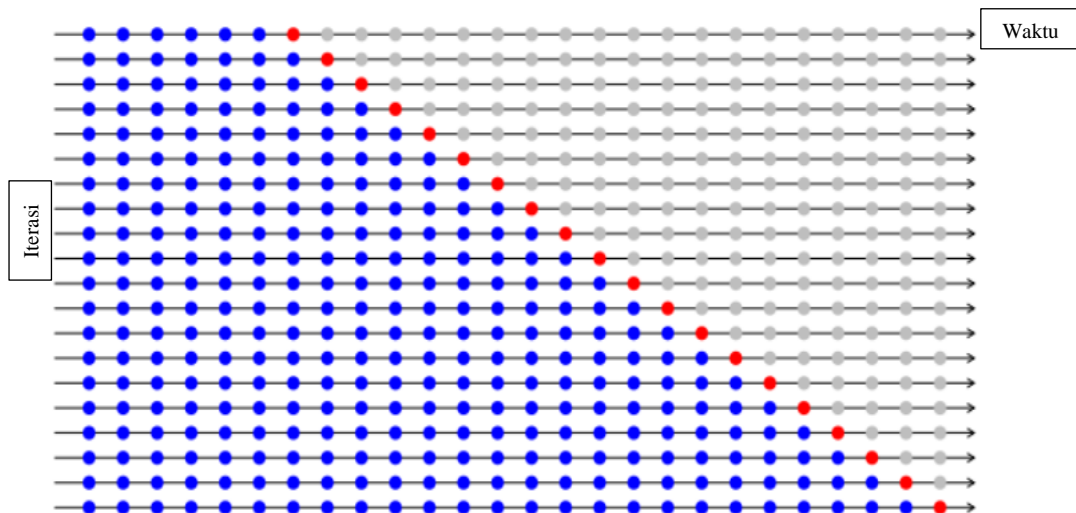


Gambar 2. 1 Ilustrasi Pembagian Data *Training* dan Data *Test* (Hyndman dan Athanasopoulos, 2018)

Komposisi pada Gambar 2.1 dibedakan antara warna biru sebagai data *training* dan warna merah sebagai data *test* dengan masing - masing bagian 80% untuk data *training* dan 20% data *test*. Namun dalam praktiknya tidak menutup kemungkinan bisa berbeda pada setiap model tergantung pada seberapa panjang data *sample* dan seberapa jauh prediksi yang diinginkan.

Teknik yang lebih canggih untuk melakukan *training* dan *testing* biasa adalah dengan cara *cross-validation*. *Time Series Cross-Validation* dipercaya mampu meminimalkan error sehingga bisa meningkatkan tingkat akurasi dari metode asli. Prosedur yang dilakukan dalam *Time Series Cross-Validation* adalah setiap kali melakukan prediksi, model *forecasting* akan diberikan kembali data *train* yang berasal dari data *test* yang sudah digunakan untuk validasi error. Sehingga prediksi berikutnya akan selalu dinamis, sudah mencakup seluruh data

aktual terkini yang tersedia maka dari proses keseluruhan dari prosedur tersebut. Proses ini biasa disebut “*Rolling Forecast*” (Hyndman dan Athanasopoulos,



Gambar 2. 2 Alur Bagaimana Data Test menjadi data Train pada Iterasi Selanjutnya (Hyndman dan Athanasopoulos, 2018)

2018). *Rolling Forecast* diilustrasikan pada Gambar 2.2.

Dari kiri ke kanan merupakan keseluruhan data 100%. Titik warna biru merupakan data *training* yang disusul 1 titik warna merah sebagai data *test*. Warna abu-abu adalah data *test* yang belum digunakan untuk pengesanan. Baris ke dua dan seterusnya pada gambar menunjukkan iterasi ke- $n$  dimana yang semula mula point berwarna merah (data *test*) berubah menjadi data *train* dan berulang hingga data *test* warna abu-abu habis.

#### 2.2.4 ARIMA

*ARIMA* (*Autoregressive Integrated Moving Average*) memodelkan dan memprediksi data dengan pola tren yang terdiri dari tiga komponen utama yang bisa disesuaikan yaitu *autoregressive* (*AR*), *moving average* (*MA*), dan *integrated* (*I*). *Komponen autoregressive* (*AR*) mengacu pada penggunaan nilai-nilai sebelumnya dalam deret waktu untuk memprediksi nilai pada waktu tertentu. *Komponen moving average* (*MA*) melibatkan penggunaan nilai-nilai sebelumnya dari kesalahan prediksi dalam deret waktu untuk memprediksi nilai pada waktu tertentu. *Komponen integrated* (*I*) melibatkan diferensiasi data untuk membuatnya stasioner, yang berarti memiliki sifat statistik yang konsisten dari waktu ke waktu.

Metode *ARIMA* diidentifikasi dengan tiga parameter utama:  $p$ ,  $d$ , dan  $q$ . Parameter  $p$  (*order of autoregressive term*) menunjukkan berapa banyak nilai sebelumnya yang akan digunakan dalam model *AR*. Parameter  $d$  (*order of differencing*) menunjukkan berapa kali deret waktu harus diferensiasi agar menjadi stasioner. Parameter  $q$  (*order of moving average term*) menunjukkan berapa banyak nilai sebelumnya dari kesalahan prediksi yang akan digunakan dalam model *MA*. Formula *ARIMA* ( $p, d, q$ ) untuk memodelkan dan memprediksi dituliskan pada persamaan 2.2.

$$ARIMA(p, d, q) = AR(p) + I(d) + MA(q) \quad (2.2)$$

Dengan :

$AR(p)$  : Komponen yang mencakup pengaruh nilai-nilai sebelumnya dalam data time series terhadap nilai di waktu sekarang.

$I(d)$  : komponen yang berhubungan dengan differensiasi data untuk membuatnya stasioner.

$MA(q)$  : Komponen *MA* adalah komponen yang mencakup pengaruh gangguan (error) pada nilai waktu sekarang berdasarkan gangguan-gangguan sebelumnya dalam data.

Orde ( $p$ ) pada  $AR(p)$  mengindikasikan berapa banyak observasi sebelumnya yang akan diperhitungkan dalam model. Model ini mengasumsikan bahwa nilai pada waktu ( $t$ ) dipengaruhi oleh nilai-nilai pada waktu ( $t - 1, t - 2$ ) hingga ( $t - p$ ). Stasioneritas adalah sifat data time series dimana statistik dasar seperti rata-rata dan varians tidak bervariasi seiring waktu. Orde ( $d$ ) pada  $I(d)$  menunjukkan seberapa banyak diferensiasi yang diterapkan pada data untuk mencapai stasioneritas. Jika ( $d$ )=0 dianggap stasioner. Jika ( $d$ )>0 itu berarti data mengalami diferensiasi sebanyak ( $d$ ) kali sebelum menjadi stasioner. . Orde ( $q$ ) mengindikasikan berapa banyak nilai gangguan sebelumnya yang akan diperhitungkan dalam model. Model  $MA(q)$  mengasumsikan bahwa nilai pada waktu ( $t$ ) dipengaruhi oleh nilai gangguan pada waktu ( $t - 1, t - 2$ ) hingga ( $t - p$ ).

### 2.2.5 *FbProphet*

Alat *open source* dalam *library python* dan R oleh Facebook ini menggabungkan keunggulan dari metode prediksi *time series* klasik dan teknik *machine learning* modern (Taylor dan Letham, 2018). Fitur utama *FbProphet* meliputi penanganan musiman ganda, tren non-linear, dan hari libur menggunakan persamaan 2.3.

$$y_t = g_t + s_t + e_t \quad (2.3)$$

dengan:

$y_t$  : Mewakili deret waktu

$g_t$  : Tren saturasi nonlinier yang memodelkan perubahan seri non periodik

$s_t$  : Komponen perubahan periodik tahunan

$e_t$  : Perubahan yang tidak biasa yang tidak dapat diakomodasi oleh model (Dash dkk., 2021).

Alat ini dipilih sebagai pembanding karena mewakili metode *machine learning* yang mudah digunakan dan fleksibel, dengan kemampuan penyesuaian musiman bisa kustom maupun otomatis sehingga tinggal pakai saja tidak banyak penyetelan. Studi ini dilakukan tuning pada interval *forecasting*-nya saja. *Network traffic LTE* yang memiliki karakteristik yang unik setiap lokasi *cell* yang berbeda akan menjadi tantangan tersendiri bagi penggunaan *FbProphet*.

### 2.2.6 *RMSE* dan *MAPE*

Kesalahan atau error dalam melakukan *forecasting* dihitung dari selisih nilai data sample asli dengan hasil prediksi. Yang dimaksudkan kesalahan ini bukanlah kesalahan model dalam melakukan hasil prediksi, namun merupakan bagian dari pengamatan yang tidak bisa diprediksi oleh model. Perbedaan kesalahan dengan residual adalah residu dihitung pada data *training* dan hanya dilihat berdasarkan satu langkah prediksi, sedangkan kesalahan disini dihitung pada data *test* dan dihitung berdasarkan beberapa langkah prediksi.

Kesalahan prediksi selain bisa dilihat dari intuisi berdasarkan visualisasinya, juga bisa dihitung secara kuantitatif dengan skala yang sama dengan besaran satuan data sample yang digunakan. *Root Mean Squared Error* (RMSE) sering

digunakan untuk mengukur kesalahan dengan angka mutlak bernilai positif dalam satuan dan skala yang sama dengan data sample. Semakin kecil *RMSE* maka diartikan semakin baik model yang digunakan. Penelitian akan menggunakan data *traffic LTE* dengan satuan Mbps, satuan *RMSE* yang dihasilkan juga Mbps. Perhitungan *RMSE* menggunakan persamaan 2.4.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad (2.4)$$

dengan:

$A_t$  = Nilai Aktual

$F_t$  = Nilai hasil prediksi

$n$  = Banyaknya data

$\sum$  = Jumlah keseluruhan nilai.

Menghitung *RMSE* adalah mengurangi nilai aktual dengan nilai hasil prediksi kemudian dikuadratkan dan dijumlahkan keseluruhan hasilnya kemudian dibagi dengan banyaknya data. Hasil perhitungan tersebut selanjutnya dihitung kembali untuk mencari nilai dari akar kuadrat. *RMSE* dipilih karena dirasa relevan dengan object penelitian yang mana memiliki kelebihan yang mampu menghitung akurasi prediksi dengan angka mutlak positif dalam satuan dan skala yang sama dengan data sample tersebut.

Selain menghitung kesalahan prediksi menggunakan *RMSE*, pekerjaan ini juga menghitung kesalahan hasil prediksi untuk menentukan seberapa layak model yang diusulkan. Uji kelayakan ini menggunakan perhitungan *Mean Absolute Percentage Error (MAPE)*. *MAPE* menghitung rata – rata perbedaan nilai absolut antara hasil prediksi dikurangi data asli kemudian dibagi nilai aslinya. Perhitungan ini dilakukan pada setiap periode observasi dengan hasil berupa prosentase (Khair dkk., 2017). Persamaan *MAPE* dalam menghitung prosentase kesalahan model dalam melakukan prediksi dijelaskan pada persamaan 2.5.

$$MAPE = \frac{\sum |A_t - F_t|}{n} \times 100\% \quad (2.5)$$

dengan:

$A_t$  = Nilai Aktual

$F_t$  = Nilai hasil prediksi

$n$  = Banyaknya data

$\sum$  = Jumlah keseluruhan nilai.

Kelebihan *MAPE* adalah hasil perhitungannya dapat diukur untuk menentukan tingkat kelayakan model dalam satuan persen. sehingga mudah diinterpretasikan dan dimengerti oleh pengambil keputusan. Ini mewakili deviasi persentase rata-rata antara nilai aktual dan nilai ramalan, memberikan indikasi yang jelas tentang besarnya kesalahan dalam peramalan. Kelebihan lainnya adalah:

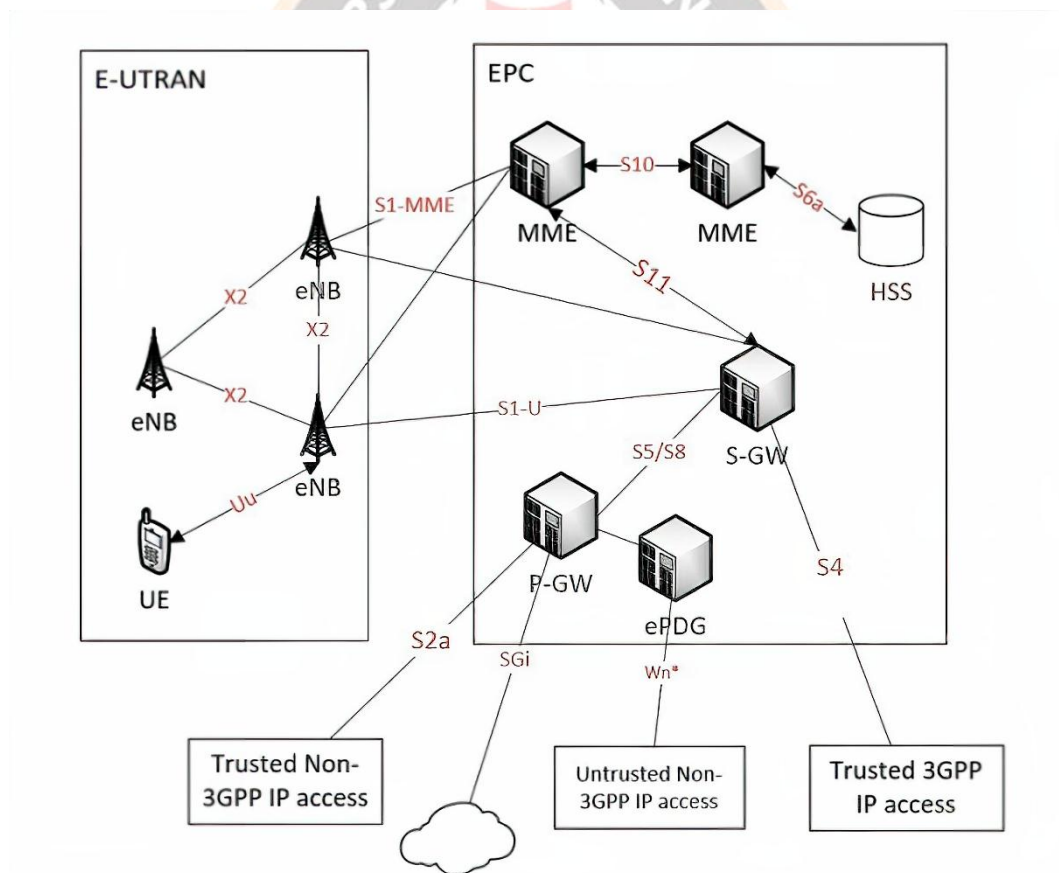
1. Ukuran Relatif: *MAPE* memberikan ukuran relatif dari akurasi peramalan, yang berguna untuk membandingkan kinerja berbagai model atau metode peramalan. Hal ini memungkinkan evaluasi standar diberbagai dataset dan periode waktu.
2. Skala-Invarian: *MAPE* tidak terpengaruh oleh skala data, sehingga dapat digunakan untuk mengevaluasi akurasi peramalan tanpa memperhatikan apakah nilai datanya besar atau kecil.
3. Penanganan Nilai Nol: *MAPE* mengatasi nilai nol dalam data aktual dengan efektif. Ketika ada nilai nol dalam data aktual, *MAPE* tidak mengalami masalah pembagian dengan nol selama nilai peramalan juga nol.
4. Digunakan Secara Luas: *MAPE* telah banyak diadopsi dalam literatur peramalan dan umum digunakan diberbagai industri dan disiplin ilmu. Penggunaan yang luas memfasilitasi perbandingan, pengukuran kinerja, dan kemampuan untuk mendapatkan wawasan dari kinerja peramalan masa lalu.

Ini sangat membantu untuk melihat tingkat akurasi metode satu dibandingkan dengan metode lainnya. Namun, perlu diingat bahwa *MAPE* juga memiliki batasan, seperti sensitivitas terhadap nilai ekstrem dan potensi masalah interpretasi ketika nilai aktual mendekati nol. Oleh karena itu, disarankan untuk

menggunakan *MAPE* bersama dengan metrik kinerja lainnya untuk memperoleh pemahaman yang komprehensif tentang akurasi peramalan.

### 2.2.7 Long Term Evolution (LTE)

*Long Term Evolution (LTE)* merupakan teknologi dalam dunia telekomunikasi yang mengacu pada standar 3GPP spesifikasi rilisan ke 8 (Jimaa dkk., 2011). Arsitektur *LTE* dikenal sebagai *System Architecture Evolution (SAE)* dimana teknologi ini adalah kebaruan generasi ke 4 dari perkembangan generasi yang ada sebelumnya yaitu *GSM* dan *UMTS* (Jimaa dkk., 2011), (Wahyudin dkk., 2017). Ada tiga komponen penting dalam arsitektur ini, *User Equipment (UE)*, *Evolved UMTS Terrestrial Radio Access Network (E-UTRAN)* dan *Evolved Packet Core (EPC)*. Gambar arsitektur *LTE* bisa dilihat pada Gambar 2.3.



Gambar 2. 3 Arsitektur *LTE* (Putra dan Widhi, 2017)

Buku "4G LTE Advanced For Beginner & Consultant" menjelaskan bahwa setiap elemen dalam Gambar 3.2 memiliki fungsi yang saling berkaitan. Pertama, *User Equipment (UE)* merupakan perangkat yang langsung berhubungan dengan pengguna, seperti modem, handphone, dan sejenisnya. Fungsi *UE* meliputi modulasi/demodulasi, transmitter, serta autentikasi dengan Terminal Equipment dan *USIM*.

Selanjutnya, *Evolved UMTS Terrestrial Radio Access Network (E-UTRAN)* menjadi elemen yang mengandung E-Node B, berfungsi sebagai transceiver untuk mengelola Radio Resource di beberapa sel. Kemudian, *Evolved UMTS Terrestrial Radio Access Network (EPC)* menjadi elemen yang paling kompleks dengan berbagai perangkat di dalamnya.

Dalam *EPC*, *Mobility Management Entity (MME)* berperan sebagai pengontrol node pada semua akses jaringan *LTE*, mengatur bearer dan signaling antara *UE* dan *Core*, serta sebagai penghubung dan *UE authenticator* untuk memilih *P-GW* atau *S-GW* untuk terhubung ke teknologi lain. *Serving Gateway (S-GW)* mengatur jalur packet dengan data dari setiap pengguna, menghubungkan *UE* dengan *EnodeB*, dan *3GPP* lainnya saat *interhandover* terjadi.

*Home Subscriber Server (HSS)* berfungsi sebagai tempat penyimpanan data pelanggan dari provider telekomunikasi. *Packet Data Network Gateway (P-GW)* menyediakan *IP Address* dari *User* ke Jaringan *packet*, menjadi pusat link antara teknologi *LTE* dengan teknologi lainnya. Terakhir, *Policy and Charging Rules (PCRF)* memiliki tugas menghitung *billing user* dari *routing* dan *charging*. Dengan demikian, setiap elemen dalam gambar tersebut memiliki peran masing-masing dalam memastikan fungsi dan kinerja jaringan *LTE* secara menyeluruh.

#### **2.2.8 Korelasi Volume Traffic, Physical Resource Block dan Throughput**

##### ***LTE***

*Cell congestion* biasanya disebabkan oleh *Physical Resource Block (PRB)* yang terpakai mencapai 100%. *Physical Resource Block (PRB)* dalam teknologi *LTE (Long-Term Evolution)* adalah unit dasar untuk alokasi sumber daya fisik dalam domain waktu dan frekuensi. *PRB* digunakan untuk mengirimkan data antara stasiun basis (*eNodeB*) dan perangkat pengguna (*UE*) dalam jaringan *LTE*.



Kenaikan alokasi *PRB* itu sendiri disebabkan oleh naiknya jumlah pengguna yang direpresentasikan oleh naiknya volume *mobile traffic* pada *cell* tersebut. *Mobile traffic* sendiri merupakan volume data yang dikonsumsi oleh pengguna ponsel seluler dalam bentuk transmisi data, seperti mengakses internet, mengunduh, atau mengunggah informasi. Kenaikan ini mengakibatkan menurunnya rata – rata *throughput* pada setiap pengguna (Chmieliauskas dan Gursnys, 2019). Kualitas layanan jaringan pada *cell* tersebut tidak akan maksimal dan perusahaan berpotensi kehilangan pendapatan karena kualitas internetnya yang buruk.

Perusahaan memiliki ambang batas untuk menentukan berapa persen beban maksimal yang bisa ditangani setiap *cell* untuk meminimalisir kerugian. Ambang batas ini bergantung pada *bandwidth* masing – masing *cell*. Misalnya *cell LTE* 800MHz dengan *bandwidth* 10 MHz dianggap mampu menangani beban dengan baik jika penggunaan *PRB* nya di bawah 60%, sedangkan untuk *cell LTE* 1800 dengan *Bandwidth* 20MHz ambang batasnya adalah kurang dari 80% (Chmieliauskas dan Gursnys, 2019).



SEKOLAH PASCASARJANA