

# Prediksi Kuantitas Hasil Budidaya Ikan Konsumsi Menggunakan Penerapan Metode Regresi Data

Yan Watequlis Syaifudin\*, Muhammad Syifa'ul Ikrom Al Masyriqi\*\*, Pramana Yoga Saputra\*\*\*, Triana Fatmawati\*\*\*\*

\* \*\* \*\* \* Program Studi Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang  
\*qulis@polinema.ac.id, \*\*ikramalmasyriqi@gmail.com, \*\*\*pramanay@polinema.ac.id, \*\*\*\* triana@polinema.ac.id

## ABSTRACT

*In the present era, the industrial 4.0 revolution marks technological advancements in the fisheries sector, making fish farming industries increasingly effective and efficient. Predictions related to future fish harvest production can be used as considerations in planning. This research aims to build a prediction system for fish farming harvests by comparing four regression methods: linear regression, polynomial regression, random forest regression, and support vector regression to obtain accurate predictions. With the existence of this harvest prediction system, it is expected to provide information about the quantity of future harvests, enabling harvest partners to take appropriate steps to enhance their profitability. Based on the research results, the best regression model is polynomial regression. This model yields an average Root Mean Square Error (RMSE) value of 4.39 and an average Mean Absolute Percentage Error (MAPE) value of 0.41%. This indicates that the regression model has good and accurate prediction capabilities.*

**Keyword:** Prediction System, Aquaculture, Harvest Partner, Regression Model

## 1. Pendahuluan

Pada zaman sekarang, revolusi industri 4.0 menandai kemajuan teknologi di sektor perikanan, yang membuat industri budidaya perikanan semakin efektif dan efisien. Selain itu, perikanan budidaya di Indonesia juga merupakan salah satu sumber ketahanan pangan. KKP (Kementerian Kelautan dan Perikanan) secara optimis mengemukakan bahwa sektor perikanan budidaya adalah sektor potensial untuk melakukan akselerasi pertumbuhan ekonomi di tahun 2022. Sebab saat ini KKP sedang mengembangkan dua terobosan baru untuk meningkatkan produktivitas perikanan budidaya. Terobosan tersebut adalah pengembangan budidaya perikanan yang berbasis ekspor dengan komoditas unggulan dan pembangunan kampung perikanan yang berbasis kearifan lokal [1]. Pernyataan KKP tersebut menunjukkan bahwa sektor perikanan budidaya dapat menjadi sebuah peluang bisnis bagi mitra panen (pengusaha budidaya perikanan).

Sektor budidaya perikanan dianggap sebagai sektor yang masih dapat tumbuh dan berkembang, karena potensi yang ada masih belum terpakai secara maksimal dibandingkan dengan luas lahan yang tersedia untuk budidaya perikanan. Karakteristik budidaya perikanan, yaitu dapat dilakukan oleh segala lapisan masyarakat, dari desa hingga kota, memberikan margin keuntungan yang cepat dan tinggi, menjadi dasar untuk meningkatkan pembangunan industri, serta dapat menerapkan teknologi yang beragam [2]. Indonesia mempunyai potensi lahan perikanan budidaya yang sangat luas [3]. Dari potensi tersebut maka diperlukan pengoptimalan budidaya perikanan oleh mitra panen untuk mencegah ketidakseimbangan ekosistem dengan mempelajari sifat hidup dan habitat asli masing-masing organisme agar perkembangan organisme atau teknik pemeliharaan yang dilakukan bisa dimanipulasi pada lingkungan budidayanya. Peningkatan produksi perikanan melalui budidaya perikanan adalah salah satu cara peningkatan produksi yang berkelanjutan dan memperhatikan kelestarian lingkungan. Populasi ikan yang ada semakin berkurang karena kegiatan penangkapan yang dilakukan secara terus-menerus khususnya di perairan umum sehingga diperlukan adanya *restocking*. Untuk melakukan *restocking*, mitra panen perlu melakukan perhitungan perkiraan permintaan konsumen dan hasil panen produksi perikanan budidaya pada waktu yang akan datang [4]. Harga jual perikanan budidaya yang saat ini terus mengalami fluktuasi dipengaruhi oleh faktor musim panen sehingga berkaitan dengan jumlah ketersediaan stok di pasar dimana hukum *supply* dan *demand* berlaku. Hal tersebut tidak terjadi pada harga pakan yang cenderung terus naik setiap tahunnya sehingga mengakibatkan terus meningkatnya biaya operasional yang harus dikeluarkan oleh mitra panen [5].

Prediksi kejadian yang terkait produksi hasil panen perikanan di masa depan dapat digunakan sebagai pertimbangan dalam perencanaan. Ilmu atau metode untuk memperkirakan atau memprediksi peristiwa di masa depan disebut *forecasting*, yang dilakukan dengan menggunakan data *time series* dan model matematis. Selain

metode *forecasting* konvensional, teknologi terbaru juga memungkinkan *forecasting* dilakukan dengan komputer dan perangkat lunak canggih, salah satunya adalah metode *machine learning* [6]. Membangun sistem prediksi budidaya perikanan yang akurat dan otomatis menjadi perhatian para pelaku industri bahkan peneliti. Pendekatan prakiraan perikanan yang ada terutama bergantung pada model prediksi umum yang biasa digunakan [7].

Penelitian kali ini akan membangun sistem prediksi hasil panen budidaya perikanan yang dapat meningkatkan profitabilitas mitra panen dengan membandingkan 4 metode regresi yaitu *linear regression*, *polynomial regression*, *random forest regression*, dan *support vector regression* guna mendapatkan prediksi yang akurat. Dengan adanya sistem prediksi hasil panen ini dapat memberikan informasi tentang jumlah hasil panen yang akan dihasilkan di masa depan, sehingga mitra panen dapat mengambil langkah-langkah yang tepat untuk meningkatkan profitabilitas mereka. Dengan mengetahui hasil panen yang diperkirakan, mitra panen dapat menyesuaikan strategi produksi untuk memaksimalkan keuntungan.

## 2. Dasar Teori

Pada bagian ini, beberapa teori yang berkaitan dengan topik yang sedang dibahas akan dibahas untuk memberikan pemahaman yang lebih komprehensif dan mendalam tentang topik tersebut.

### 2.1. Perikanan Budidaya

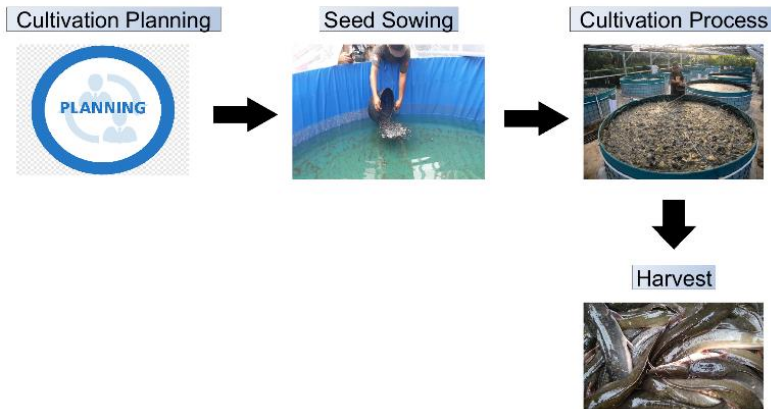
Perikanan budidaya merupakan hasil produksi ikan yang diperoleh melalui cara budidaya di ladang, di sawah atau ditempat lain yang berada di daratan. Ada beberapa definisi terkait perikanan budidaya antara lain: Menurut Badan Pusat Statistik Kabupaten Malang (2019), perikanan budidaya diklasifikasikan atas jenis budidaya yaitu budidaya laut, tambak, kolam. Karamba, jarring apung dan sawah.

Berdasar Undang-undang Republik Indonesia Nomor 31 Tahun 2004 pembudidayaan ikan merupakan aktivitas untuk memelihara, membesarkan, dan/atau membiakkan ikan serta memanen hasilnya dalam lingkungan yang terkontrol, termasuk aktivitas yang menggunakan kapal untuk memuat, mengangkut, menyimpan, mendinginkan, menangani, memasak, dan atau mengawetkannya. Perikanan budidaya adalah budidaya organisme air, termasuk ikan, moluska, kurtasea dan flora air yang mencakup beberapa bentuk kegiatan dalam proses pemeliharaan untuk menaikkan produksi, seperti penebaran yang teratur, hadiah kuliner/pakan, perlidungan berdasarkan predator dan lain-lain [8].

### 2.2. Proses Budidaya Ikan

Proses budidaya ikan dimulai dengan perencanaan yang matang. Pada tahap ini, pembudidaya melakukan estimasi jumlah pakan dan benih yang akan menjadi target panen selama proses budidaya. Untuk mencapai hal ini, dilakukan pencatatan jumlah benih yang akan ditebar, berat benih dalam gram, dan perhitungan kebutuhan pakan yang diperlukan untuk setiap kolam. Ukuran kolam yang digunakan ditentukan berdasarkan jumlah benih yang akan ditebar. Sebagai contoh, untuk 1500 ekor bibit ikan lele, digunakan kolam dengan diameter 3 meter dan tinggi 1 meter, sehingga volume kolam yang dibutuhkan adalah sekitar 7 meter kubik [9]. Selanjutnya, proses budidaya dilanjutkan dengan melakukan pengadaan benih. Benih ikan yang dibesarkan berasal dari induk ikan berkualitas, yang diperoleh dari pembudidaya benih yang memiliki reputasi baik dalam pemuliaan induk ikan berkualitas.

Setelah proses pengadaan pakan dan benih selesai ditebarkan pada kolam tahap selanjutnya adalah proses budidaya pembesaran yang berlangsung kurang lebih 90 hari untuk ikan lele tergantung ukuran benih dan target panen yang ditentukan. Proses pemberian pakan dilakukan dengan tetap melihat kondisi cuaca dan kesehatan ikan, bisa 2 hingga 3 kali sehari hal ini juga mempengaruhi kecepatan pertumbuhan ikan dengan target ukuran 8-10 ekor/kg sehingga rata-rata berat ikan per ekor ketika panen sekitar 100-125 gram. Dalam proses budidaya ini pembudidaya melakukan proses sorting pada ikan setiap kurang lebih 25 hari sekali untuk memisahkan ikan yang memiliki tingkat pertumbuhan lebih tinggi daripada yang lainnya karena sifat kanibalisme yang dimiliki oleh jenis ikan lele dapat merugikan pembudidaya. Setelah masa panen tiba pembudidaya memiliki beberapa opsi penjualan melalui pedagang besar, *reseller* maupun langsung kepada *end user* dengan tingkat keuntungan yang berbeda-beda [5].



Gambar 1. Proses Budidaya Ikan

### 2.3. *Survival Rate*

Tingkat kelangsungan hidup (*Survival Rate*) adalah rasio jumlah ikan yang berhasil bertahan hidup dari awal hingga akhir proses budidaya. Tingkat kelangsungan hidup atau *Survival Rate* memiliki dampak signifikan terhadap hasil produksi yang diperoleh, di mana semakin tinggi *Survival Rate*, semakin besar hasil panen yang dapat dicapai oleh pembudidaya. Adanya nilai *Survival Rate* yang rendah diduga disebabkan oleh kurangnya penerapan *Quality Control* dan aklimatisasi yang tepat saat melakukan penyebaran benih ikan. Hal ini mengakibatkan ikan tidak dapat beradaptasi dengan baik dan akhirnya mengalami kematian [10]. Untuk menghitung tingkat kelangsungan hidup, digunakan rumus sebagai berikut [11].

$$SR = \frac{Nt}{NO} \times 100\%$$

Keterangan:

SR: *Survival rate* (%)

Nt: Jumlah ikan pada akhir budidaya

NO: Jumlah ikan pada awal budidaya

### 2.4. Sistem Prediksi

Sistem merupakan kesatuan unsur atau komponen yang saling berhubungan untuk memudahkan perpindahan materi, energi, atau informasi. Sistem ini disebut juga sebagai satu komponen yang saling berhubungan dan memiliki elemen penggerak. Secara umum unsur-unsur yang membentuk suatu sistem dan perlu diketahui adalah tujuan, input, proses, dan output. Ada juga faktor lain seperti batasan, mekanisme kontrol dan umpan balik, lingkungan sistem, dan lainnya. Prediksi adalah proses memprediksi variabel masa depan yang lebih intuitif daripada data historis, tetapi lebih intuitif, data kuantitatif sering digunakan dalam peramalan harga sebagai pelengkap peramalan. Maka dapat disimpulkan sistem prediksi adalah suatu kesatuan unsur atau komponen yang memiliki tujuan untuk dapat memproses variabel tertentu yang digunakan untuk memprediksi variabel baru berdasarkan variabel yang telah diolah atau ada [12].

### 2.5. *Multiple Linear Regression (MLR)*

MLR atau Regresi Linier Berganda merupakan teknik untuk menganalisis hubungan antara 2 variabel yaitu, variabel dependen ( $y$ ) terhadap variabel independent ( $x_1, x_2, x_3, \dots, x_n$ ) secara linier sehingga terbentuknya suatu model. Tujuan metode *Multiple Linear Regression (MLR)* ialah untuk menghasilkan suatu model pada variabel dependen ( $y$ ) berdasarkan nilai-nilai dari variabel independen ( $x_1, x_2, x_3, \dots, x_n$ ) [13].

### 2.6. *Support Vector Regression*

*Support Vector Regression (SVR)* merupakan suatu metode SVM yang diterapkan pada kasus regresi. SVR bertujuan untuk menemukan sebuah fungsi  $f(x)$  sebagai suatu hyperplane (garis pemisah) berupa fungsi regresi yang mana sesuai dengan semua masukan data dengan membuat error ( $\epsilon$ ) sekecil mungkin. Metode SVR merupakan metode yang dapat menghasilkan performansi yang baik karena dapat mengatasi masalah overfitting, yaitu sebuah kondisi dimana model pembelajaran mesin memiliki error yang sangat rendah pada data pelatihan tetapi sangat tinggi pada data pengujian, dan jika didapatkan nilai  $\epsilon=0$  artinya didapatkan suatu regresi yang sempurna [14].

### 2.7. *Polynomial Regression*

Regresi polinomial adalah jenis regresi khusus yang bekerja pada hubungan lengkung antara nilai dependen dan nilai independen. Regresi Polinomial merupakan model Regresi Linear yang dibentuk dengan menjumlahkan pengaruh masing-masing variabel prediktor (X) yang dipangkatkan meningkat sampai orde ke-n. Regresi polinomial merupakan hasil modifikasi dari model regresi linier berganda. Untuk membedakan dengan regresi linear berganda, secara umum regresi polinomial dapat dinobatkan dalam bentuk berpangkat. Selanjutnya proses mencari turunan parsial dari SSE terhadap koefisien beta dan menyamadengankan dengan nol seperti pada regresi linier berganda [15].

### 2.8. Random Forest Regression

*Random forest regression* merupakan algoritma yang berasal dari kombinasi beberapa *decision tree*. Setiap *decision tree* pada *random forest* bergantung pada nilai suatu vektor acak yang sampelnya diambil secara independen dan distribusi yang sama. Pada algoritma *random forest*, parameter jumlah estimator menunjukkan banyaknya *decision tree* dalam *forest*. Secara umum, semakin banyak jumlah estimator, semakin baik hasilnya. Namun, pada suatu titik tertentu kinerja prediksi akan menurun dikarenakan tingginya kebutuhan komputasi. Oleh karena itu perlu dicari di jumlah estimator berapa algoritma menghasilkan nilai prediksi terbaik [16].

### 2.9. Laravel

Laravel merupakan salah satu framework PHP terbaik saat ini yang dikembangkan oleh Taylor Otwell yang hadir sebagai platform *web development* bersifat *open source*. Laravel memiliki sintaks yang dirancang guna memudahkan dan mempercepat proses pengembangan *website* dengan gaya yang ekspresif dan elegan. Meskipun Laravel bukan satu-satunya *framework* PHP yang banyak digunakan, akan tetapi bisa menjadi opsi yang dapat dipertimbangkan karena bersifat *open source* dan banyak fitur yang dapat digunakan [17].

### 2.10. MySQL

MySQL merupakan program pengakses database yang bersifat jaringan sehingga dapat digunakan untuk aplikasi multi user. Kelebihan lain dari MySQL adalah MySQL menggunakan bahasa query standar yang dimiliki SQL. Kepopuleran MySQL antara lain karena MySQL menggunakan Sql sebagai bahasa dasar untuk mengakses Database-nya sehingga mudah untuk digunakan. MySQL juga bersifat *open source* dan *free* pada berbagai platform kecuali pada windows yang bersifat *shareware*. MySQL didistribusikan dengan lisensi *open source* GPL (*General Public License*) mulai versi 3.23, pada bulan Juni 2000 [18].

## 3. Metode Penelitian

Bagian ini membahas desain sistem prediksi hasil panen dan proses pengembangan aplikasi. Bagian ini menjelaskan langkah-langkah yang diambil untuk membuat sistem prediksi hasil panen, serta beberapa faktor penting yang memengaruhi hasil panen dan pengembangan aplikasi.

### 3.1. Pengumpulan Data

Pada penelitian ini, subjek yang ditargetkan adalah pengusaha budidaya perikanan yang merupakan mitra panen dari PT Adma Digital Solusi. Metode yang digunakan untuk mengumpulkan data berasal dari kolam yang dibudidaya oleh mitra panen, yang meliputi data pembenihan, monitor pemberian pakan, monitor kematian budidaya, rata-rata berat budidaya, dan hasil panen budidaya.

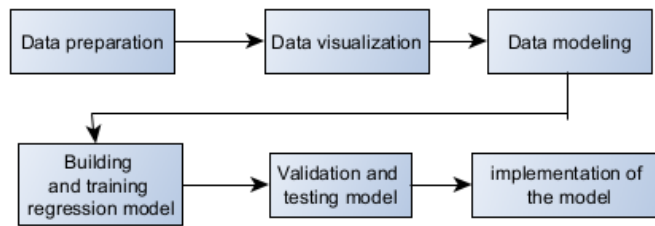
### 3.2. Faktor Budidaya Perikanan

Berdasarkan studi literatur mengenai budidaya ikan dan data-data yang telah dikumpulkan, teridentifikasi beberapa faktor penting dalam budidaya ikan yang dapat digunakan untuk membangun sebuah sistem prediksi panen perikanan. Faktor-faktor tersebut antara lain:

1. Jumlah bibit yang akan dibudidayakan.
2. Berat rata-rata bibit per ekor dalam gram.
3. Volume kolam yang digunakan untuk budidaya dalam meter kubik.
4. Berat rata-rata ikan per ekor saat panen (yang diperoleh dari 1kg dibagi dengan jumlah ikan yang akan dijual).
5. Persentase tingkat kelangsungan hidup ikan.
6. Total pakan yang diberikan selama budidaya dalam kilogram.
7. Hasil panen dalam kilogram.

### 3.3. Pengolahan Data

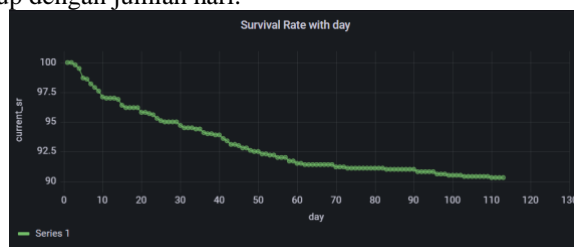
Data yang dikumpulkan perlu melalui tahap pengolahan data agar dapat mengimplementasikan model regresi. Langkah-langkah pengolahan data dapat dilihat pada diagram blok berikut.



Gambar 2. Diagram Blok Pengolahan Data

Dari diagram blok alur proses pengolahan data di atas, terdapat beberapa tahapan dalam pengolahan data pada penelitian ini dengan proses sebagai berikut:

1. **Data Preparation:** pada proses ini dilakukan prosedur pembersihan data, seperti menghilangkan data yang tidak relevan, dan menangani data yang kosong.
2. **Data Visualization:** pada proses ini data yang telah disiapkan sebelumnya divisualisasikan untuk melihat hubungan antar data. Berikut ini adalah contoh visualisasi data persentase tingkat kelangsungan hidup dengan jumlah hari.



Gambar 3. Visualisasi Data Tingkat Kelangsungan Hidup dengan Jumlah Hari

3. **Data Modeling:** setelah mengetahui hubungan antar data, maka dilakukan proses pemodelan data yang nantinya akan digunakan sebagai dataset atau input untuk model regresi yang digunakan. Dari data harian sebelumnya, data tersebut selanjutnya dikategorikan berdasarkan hasil panen dan jenis ikan. Berikut ini adalah contoh dataset yang digunakan.

Tabel 1. Sample Datasets

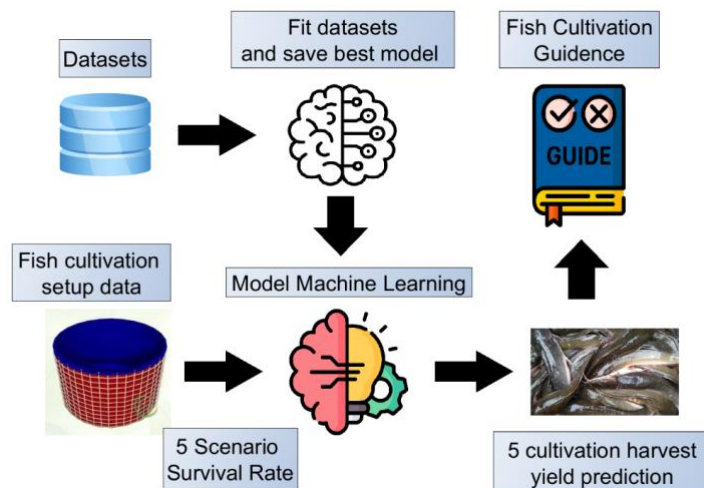
fish_id	seed_amount	seed_weight (g)	survival_rate (%)	average_weight (g)	Pond_volume (m3)	total_feed_spent (kg)
1	1000	4	90,3	115	3	167,05
1	1000	4,5	90,5	115	3	168,05
1	1500	4,2	90,1	125	7	246,45
1	2000	3,5	90,1	115	12,5	342,45
2	1000	15,5	90,4	500	7	639,85
2	1500	16,5	90,8	500	12,5	886,85
2	2000	17	90,2	500	19,5	1150,95
3	1000	16,5	72,5	500	7	453,15
3	1500	17	72,6	500	12,5	650,65
3	2000	18	81,5	500	19,5	1145,15

4. **Building and Training Regression Model:** pada proses ini membuat dan melatih 4 model regresi yaitu *multiple linear regression*, *SVR (Support Vector Regression)*, *polynomial regression*, dan *random forest regression* menggunakan data yang telah dimodelkan sebelumnya.
5. **Validation and Testing Model:** pada proses ini melakukan validasi dan uji empat model yang telah dibuat menggunakan data yang tidak digunakan dalam proses pelatihan. Pengujian model menggunakan metode MAPE (*Mean Absolute Percentage Error*) dan RMSE (*Root Mean Square Error*).
6. **Implementation of The Model:** setelah membangun, melatih, memvalidasi, dan menguji model regresi, langkah selanjutnya adalah menerapkannya pada situasi nyata. Pada proses ini, model regresi dengan akurasi terbaik diimplementasikan pada sistem untuk memprediksi hasil panen budidaya perikanan pada menu prediksi hasil panen.

### 3.4. Proses Prediksi Hasil Panen

Pada proses ini, model dengan akurasi terbaik diimplementasikan dalam sistem untuk melakukan prediksi hasil panen budidaya perikanan pada menu "Prediksi Hasil Panen" yang tersedia bagi mitra panen. Melalui menu ini, mitra panen dapat memprediksi hasil panen budidaya perikanan dalam satuan kilogram sebelum memulai budidaya. Untuk melakukan prediksi, mitra panen perlu memasukkan data persiapan awal budidaya, seperti jenis ikan, jumlah bibit dalam ekor, rata-rata berat bibit per ekor dalam gram, volume kolam dalam meter kubik, dan target rata-rata berat ikan per ekor yang akan dijual dalam gram (dihitung dengan membagi 1000g dengan jumlah ekor yang akan dijual dalam 1kg). Data persentase tingkat kelangsungan hidup ikan akan diinputkan secara otomatis oleh sistem menggunakan referensi dari data sebelumnya, yang menyediakan 5 persentase tingkat kelangsungan hidup. Hal ini akan menghasilkan 5 hasil prediksi yang disajikan kepada mitra panen. Penentuan 5 data persentase tingkat kelangsungan hidup menggunakan data minimal dan maksimal, lalu intervalnya disesuaikan untuk memperoleh 5 data input persentase tingkat kelangsungan hidup. Selanjutnya, data total pakan yang akan diberikan selama budidaya akan diinputkan secara otomatis oleh sistem, menggunakan referensi dari data sebelumnya dengan mempertimbangkan jenis ikan, persentase tingkat kelangsungan hidup, dan jumlah bibit. Dari data tersebut, akan dicari rata-rata total pakan yang diberikan.

Data-data yang telah dimasukkan akan digunakan sebagai input untuk model machine learning terbaik yang telah diuji sebelumnya. Setelah memproses data input, model akan menghasilkan output berupa 5 prediksi hasil panen budidaya dalam satuan kilogram berdasarkan 5 skenario persentase tingkat kelangsungan hidup. Sistem akan secara otomatis memberikan panduan dan simulasi pemberian pakan budidaya sesuai dengan hasil prediksi yang disajikan, seperti yang ditunjukkan pada ilustrasi berikut.



Gambar 4. Proses Prediksi Hasil Panen

## 4. Hasil dan Analisis

Pada bagian ini akan dibahas mengenai hasil dan analisis evaluasi metode regresi yang digunakan.

### 4.1. Skenario Evaluasi

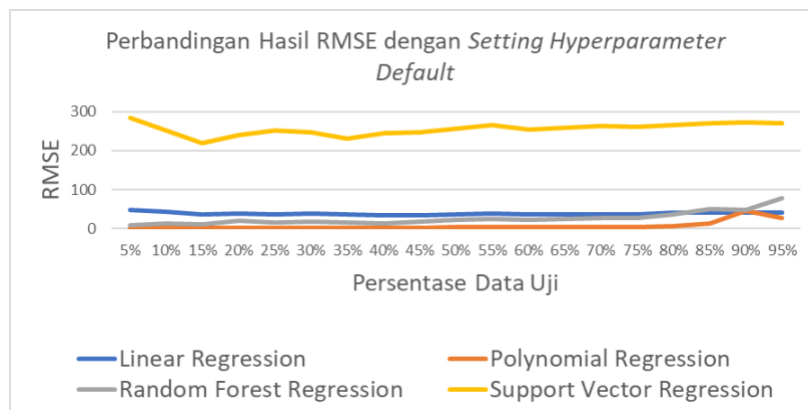
Pada penelitian ini dilakukan evaluasi terhadap metode yang diadopsi. Evaluasi metode yang diadopsi bertujuan untuk mengukur keakuratan model regresi dengan menggunakan metode RMSE (*Root Mean Square Error*) dan MAPE (*Mean Absolute Percentage Error*). Pengujian ini bertujuan untuk memahami pengaruh persentase data uji dan pengaturan hyperparameter terhadap kinerja empat model regresi yang berbeda. Skenario pengujian melibatkan penggunaan 19 persentase data uji yang bervariasi dari 5% hingga 95%, dengan interval yang meningkat sebesar 5%. Selain itu, dua pengaturan hyperparameter yang digunakan adalah pengaturan default dan pengaturan terbaik. Tujuan utama dari pengujian ini adalah untuk menganalisis kemampuan model regresi dalam mengenali pola data, serta mengevaluasi nilai RMSE dan MAPE yang dihasilkan.

### 4.2. Hasil Evaluasi RMSE dengan *Setting Hyperparameter Default*

Berdasarkan hasil pengujian dengan pengaturan hyperparameter default, dapat dilihat bahwa model Regresi Linier memiliki nilai RMSE terendah pada persentase data uji 45% (34.196) dan tertinggi pada persentase data uji 5% (49.032), sedangkan Regresi Polinomial memiliki nilai RMSE terendah pada persentase

data uji 5% (1.675) dan tertinggi pada persentase data uji 90% (46.37). *Random Forest Regression* mencapai nilai RMSE terendah pada persentase data uji 5% (8,182) dan tertinggi pada persentase data uji 95% (77,004), sedangkan *Support Vector Regression* mencapai nilai RMSE terendah pada persentase data uji terendah pada persentase data uji 15% (218,493) dan tertinggi pada persentase data uji 5% (283,76). persentase data uji 5% (283,761).

Secara rata-rata, model Regresi Polinomial memiliki RMSE terendah (7.170), diikuti oleh Regresi Random Forest (25.940), Regresi Linier (38.390), dan Regresi Support Vector (255.470). Nilai RMSE terendah secara keseluruhan adalah 1,675 untuk persentase data uji 5% pada model Regresi Polinomial, sedangkan nilai RMSE tertinggi adalah 283,761 pada persentase data uji 5% pada model Regresi *Support Vector*.

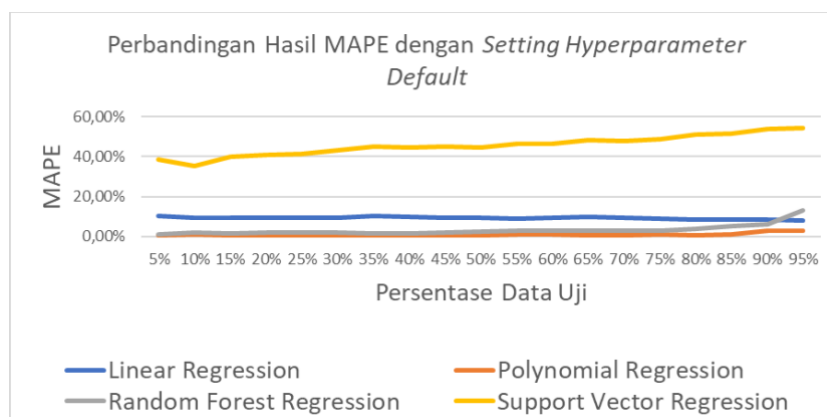


Gambar 5. Hasil Evaluasi RMSE dengan Setting Hyperparameter Default

### 4.3. Hasil Evaluasi MAPE dengan Setting Hyperparameter Default

Berdasarkan hasil pengujian dengan setting hyperparameter default, dapat dilihat bahwa model Regresi Linier memiliki nilai MAPE terendah pada persentase data uji 95% (8%) dan tertinggi pada persentase data uji 35% (10,21%), sedangkan Regresi Polinomial memiliki nilai MAPE terendah pada persentase data uji 20% (0,47%) dan tertinggi pada persentase data uji 90% (2,93%). *Random Forest Regression* mencapai nilai MAPE terendah pada persentase data uji 5% (0.91%) dan tertinggi pada persentase data uji 95% (12.87%), sedangkan *Support Vector Regression* mencapai nilai MAPE terendah pada persentase data uji terendah pada persentase data uji 10% (35.45%) dan tertinggi pada persentase data uji 95% (54.49%).

Secara rata-rata, model Regresi Polinomial memiliki nilai MAPE terendah (0.9%), diikuti oleh Regresi *Random Forest* (3.16%), Regresi Linier (9.18%), dan Regresi *Support Vector* (45.56%). Nilai MAPE terendah secara keseluruhan adalah 0.47% untuk persentase data uji 20% pada model Regresi Polinomial, sedangkan nilai MAPE tertinggi adalah 54.49% pada persentase data uji 95% pada model Regresi *Support Vector*.



Gambar 6. Hasil Evaluasi MAPE dengan Setting Hyperparameter Default

### 4.4. Hasil Setting Hyperparameter Terbaik

Pada penelitian ini, digunakan algoritma GridSearchCV untuk mengatur *hyperparameter* model regresi dan memperoleh parameter terbaik. Salah satu kelebihan utama dari GridSearchCV adalah kompatibilitasnya yang sangat baik dengan berbagai algoritma *machine learning*. Algoritma ini mampu menemukan parameter terbaik secara objektif dengan menggunakan metode *cross-validation*, di mana skor

kinerja model diperoleh pada setiap kombinasi parameter. Selain itu, GridSearchCV juga efektif dalam mencegah *overfitting* dengan menerapkan teknik validasi silang. Langkah pertama untuk menentukan *hyperparameter* terbaik yaitu dilakukan proses inialisasi parameter estimator untuk setiap model regresi seperti pada potongan kode program berikut.

Tabel 2. Cuplikan kode inisiasi parameter estimator

```

model_params = {
    'linear_regression': {
        'model': LinearRegression(),
        'params': {
            # No Parameter
        }
    },
    'polynomial_regression': {
        'model': pipeline_poly,
        'params': {
            'poly__degree': [2,3,4]
        }
    },
    'random_forest': {
        'model': RandomForestRegressor(),
        'params': {
            'n_estimators': [1,10,100],
            'min_samples_leaf': [1,2,3],
            'random_state': [0]
        }
    },
    'svr': {
        'model': pipeline_svr,
        'params': {
            'svr__kernel': ["linear", "poly", "rbf", "sigmoid", "precomputed"],
            'svr__degree': [2,3,4],
            'svr__gamma': ['scale', 'auto'],
            'svr__coef0': [0,1,2],
            'svr__C': [1,10,100],
        }
    }
}
    
```

Selanjutnya GridSearchCV akan mengeksplorasi dan mengevaluasi kombinasi parameter dari estimator, kemudian menerapkan skenario tersebut pada model untuk mencari parameter terbaik. Hasil dari pencarian parameter terbaik tersebut dapat terlihat pada Tabel 3 berikut.

Tabel 3. Hasil *Setting Hyperparameter* menggunakan GridSearchCV

No.	Model	Skor Terbaik	Parameter Terbaik
1.	Linear Regression	-32,151	Tidak memiliki parameter
2.	Polynomial Regression	-0,237	Degree: 3
3.	Random Forest Regression	-13,144	Min_samples_leaf: 1, n_estimators: 100, random_state: 0
4.	Support Vector Regression	-0,477	C: 100, coef0: 2, degree: 3, gamma: scale, kernel: poly

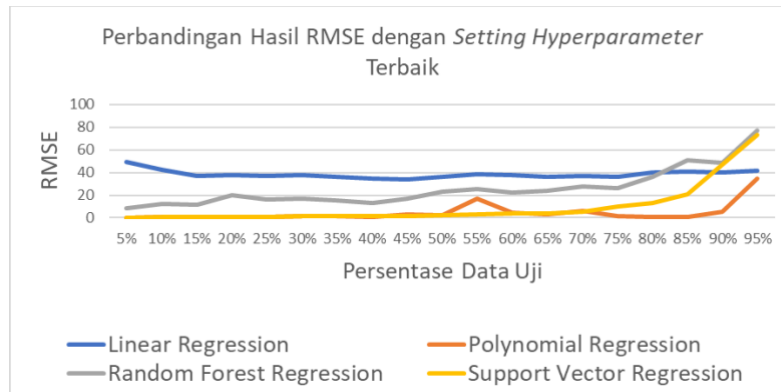
Setelah mendapatkan parameter terbaik untuk setiap model regresi dengan menggunakan GridSearchCV, maka langkah selanjutnya adalah melakukan uji coba untuk mengukur nilai RMSE dari setiap model regresi dengan menggunakan setting *hyperparameter* terbaik.

**4.5. Hasil Evaluasi RMSE dengan *Setting Hyperparameter* Terbaik**



Berdasarkan hasil pengujian dengan pengaturan hyperparameter terbaik, dapat diketahui bahwa model Regresi Linier memiliki nilai RMSE terendah pada persentase data uji 45% (34,196) dan tertinggi pada persentase data uji 5% (49,032), sedangkan model Regresi Polinomial memiliki nilai RMSE terendah pada persentase data uji 5% (0,168) dan tertinggi pada persentase data uji 95% (34,72). Regresi Random Forest mencapai nilai RMSE terendah pada persentase data uji 5% (8,182) dan tertinggi pada persentase data uji 95% (77,004), sedangkan Support Vector Regression memiliki nilai RMSE terendah pada persentase data uji 5% (9,127) dan tertinggi pada persentase data uji 95% (75,72).

Secara rata-rata, model Regresi Polinomial memiliki nilai RMSE terendah (4.39), diikuti oleh Regresi Support Vector (9.97), Regresi Random Forest (25.94), dan Regresi Linier (38.39). Nilai RMSE terendah secara keseluruhan adalah 0.15 untuk persentase 5% data uji pada model Support Vector Regression, sedangkan nilai RMSE tertinggi adalah 77.004 pada persentase 95% data uji pada model Random Forest Regression.

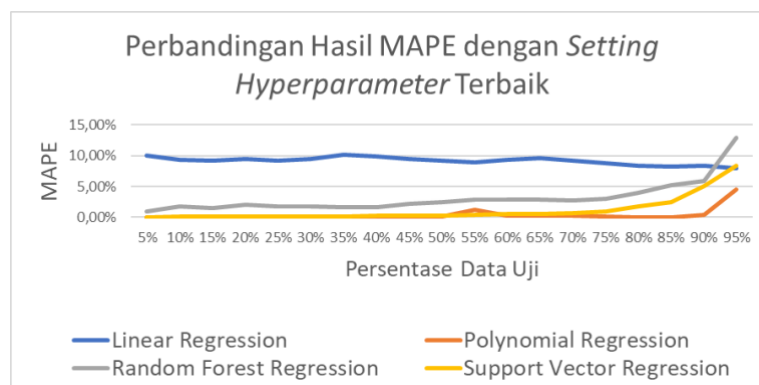


Gambar 7. Hasil Evaluasi RMSE dengan Setting Hyperparameter Terbaik

#### 4.6. Hasil Evaluasi MAPE dengan Setting Hyperparameter Terbaik

Berdasarkan hasil pengujian dengan pengaturan hyperparameter terbaik, dapat dilihat bahwa model Regresi Linier memiliki nilai MAPE terendah pada persentase data uji 95% (8%) dan tertinggi pada persentase data uji 35% (10.21%), sedangkan Regresi Polinomial memiliki nilai MAPE terendah pada persentase data uji 5% (0.02%) dan tertinggi pada persentase data uji 95% (4.46%). Regresi Random Forest mencapai nilai MAPE terendah pada persentase data uji 5% (0.91%) dan tertinggi pada persentase data uji 95% (12.87%), sedangkan Regresi Support Vector mencapai nilai MAPE terendah pada persentase data uji 5% (0.04%) dan tertinggi pada persentase data uji 95% (8.31%).

Secara rata-rata, model Regresi Polinomial memiliki nilai MAPE terendah (0.41%), diikuti oleh Regresi Support Vector (1.18%), Regresi Random Forest (3.16%), dan Regresi Linier (9.18%). Nilai MAPE terendah secara keseluruhan adalah 0.02% untuk persentase 5% data uji pada model Regresi Polinomial, sedangkan nilai MAPE tertinggi adalah 10.55% pada persentase 95% data uji pada model Regresi Random Forest.



Gambar 8. Hasil Evaluasi MAPE dengan Setting Hyperparameter Terbaik

### 5. Kesimpulan

Berdasarkan penelitian yang telah dilakukan terhadap Sistem Prediksi Hasil Panen Perikanan Budidaya untuk Meningkatkan Profitabilitas Mitra Panen, dapat disimpulkan bahwa metode regresi polinomial dengan pengaturan hyperparameter terbaik merupakan model regresi terbaik yang layak untuk digunakan pada sistem. Hasil evaluasi menunjukkan bahwa model tersebut memiliki rata-rata RMSE yaitu besarnya tingkat kesalahan

hasil prediksi yang rendah yaitu sebesar 4,39 dan rata-rata MAPE yaitu persentase kesalahan hasil peramalan yang kecil, sebesar 0,41%. Hal ini menunjukkan bahwa model regresi memiliki kemampuan prediksi yang sangat baik dan akurat.

Dengan adanya sistem ini, mitra panen memiliki alat yang berguna untuk mengoptimalkan operasi pertanian mereka dan mencapai hasil yang lebih baik secara finansial. Dengan memanfaatkan prediksi yang akurat, mereka dapat membuat keputusan yang lebih baik dalam manajemen budidaya mereka. Dengan demikian, sistem prediksi hasil panen ini memberikan kontribusi nyata dalam meningkatkan profitabilitas bisnis budidaya perikanan bagi mitra panen.

### Acknowledgements

Penelitian ini didukung oleh Jurusan Teknologi Informasi, Politeknik Negeri Malang dan PT. Adma Digital Solusi.

### Daftar Pustaka

- [1] BPS, "Statistik Sumber Daya Laut Dan Pesisir," *Badan Pus. Stat.*, p. Katalog BPS / BPS Catalogue : 3312002, 2022.
- [2] A. Noviani *et al.*, "Peramalan Produksi Budidaya Perikanan Dengan Metode Single Moving Average di UPR. Sukses Selele Kelurahan Mutiara," *JUTSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 2, no. 2, pp. 135–142, 2022.
- [3] K. Kelautan and D. A. N. Perikanan, "Lampiran I : RENCANA STRATEGIS KEMENTERIAN KELAUTAN DAN PERIKANAN I 1," pp. 1–108, 2020.
- [4] N. Permatasari, M. Hafiyusholeh, and S. Purwanto, "Forecasting Hasil Produksi Perikanan Budidaya Laut Menggunakan ARIMA," *J. Mhs. Mat. Algebr.*, vol. 1, no. 1, pp. 73–80, 2020, [Online]. Available: <https://core.ac.uk/download/pdf/329047547.pdf>
- [5] M. A. D. Ardianto, "Pengembangan Kerangka Kerja Supply Chain dan Desain Ekosistem untuk Melakukan Digitalisasi Industri Perikanan Lele di Jawa Timur," 2022, [Online]. Available: <https://repository.its.ac.id/93863/>
- [6] A. Pamungkas, L. Thesiana, and K. Adiyana, "Implementasi Peramalan dalam Industrialisasi Perikanan," *Pros. Semin. Nas. Apl. Sains Teknol.*, pp. 39–45, 2021.
- [7] J. Zhu, J. Xu, C. Guo, X. Zhuo, Y. Gao, and S. Xing, "An empirical research of marine fishery forecasting methods based on the classification model," *PIC 2016 - Proc. 2016 IEEE Int. Conf. Prog. Informatics Comput.*, no. 61402264, pp. 1–7, 2017, doi: 10.1109/PIC.2016.7949456.
- [8] F. E. Supriatin and A. N. Rohman, "Peramalan Produksi Perikanan Budidaya di Kabupaten Malang Dengan Metode Exponential Smoothing," *J. Media Akuatika*, vol. 5, no. 2, p. 51, 2020, doi: 10.33772/jma.v5i2.11961.
- [9] Y. A. D. Susanti, Z. Pramudia, A. A. Amin, L. N. Salamah, A. T. Yanuar, and A. Kurniawan, "Peningkatan Produksi Pangan melalui Sistem Integrasi Teknologi Aquaponics-Recirculating Aquaculture System (A-RAS) pada Budidaya Ikan Lele di Desa Kaliuntu Kabupaten Tuban," *Rekayasa*, vol. 14, no. 1, pp. 121–127, 2021, doi: 10.21107/rekayasa.v14i1.10254.
- [10] H. Haliyani, "PERFORMANSI KINERJA BUDIDAYA PEMBESARAN IKAN LELE SANGKURIANG (*Clarias gariepinus*) DI CV. DAMPO AWANG KECAMATAN JATINANGOR KABUPATEN SUMEDANG, JAWA BARAT," *Bul. Jalandhitah Sarva Jivitam*, vol. 2, no. 1, p. 1, 2021, doi: 10.15578/bjsj.v2i1.8445.
- [11] D. Setyani, Y. Mantuh, and T. S. Augusta, "BUDIDAYA IKAN LELE DUMBO (*Clarias gariepinus*) DAN IKAN NILA HITAM (*Oreochromis niloticus*) DALAM EMBER (BUDIKDAMBER)," *Ziraa'Ah Maj. Ilm. Pertan.*, vol. 46, no. 2, p. 157, 2021, doi: 10.31602/zmip.v46i2.4313.
- [12] H. Said, N. H. Matondang, and H. N. Irmada, "Penerapan Algoritma K-Nearest Neighbor Untuk Memprediksi Kualitas Air Yang Dapat Dikonsumsi," *Techno.Com*, vol. 21, no. 2, pp. 256–267, 2022, doi: 10.33633/tc.v21i2.5901.
- [13] W. P. Juliastoro, I. Cholissodin, and F. A. Bachtar, "Sistem Prediksi Hasil Produksi Udang Vaname menggunakan Algoritma Multiple Linear Regression ( MLR ) Kombinasi Gradient Descent ( GD ) dengan Apache Spark," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 11, pp. 4920–4928, 2021.
- [14] F. Fadillah, S. A. Wibowo, G. Budiman, F. T. Elektro, and U. Telkom, "APLIKASI BERBASIS ANDROID MENGGUNAKAN METODE SUPPORT VECTOR REGRESSION DESIGN AND IMPLEMENTATION OF STOCK PRICE PREDICTION IN ANDROID-," vol. 7, no. 2, pp. 3869–3876, 2020.
- [15] A. Eka, A. Juarna, T. Informatika, F. T. Industri, and U. Gunadarma, "Prediksi Produksi Daging Sapi Nasional dengan Metode Regresi Linier dan Regresi Polinomial Pendahuluan Regresi Linier Regresi Polinomial," vol. 20, pp. 209–215, 2021.
- [16] N. KURNIAWATI, A. NOVITRI, and Y. K. NINGSIH, "Prediksi Channel Gain Threshold untuk Modulasi Adaptif V2V menggunakan Algoritma Random Forest Regression," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 10, no. 3, p. 544, 2022, doi: 10.26760/elkomika.v10i3.544.
- [17] H. D. Budiana, S. Lestanti, and S. Nur Budiman, "Aplikasi Prediksi Stok Popok Bayi Berbasis Web Dengan Metode Trend Moment," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 6, no. 1, pp. 26–30, 2021, doi: 10.36040/jati.v6i1.4362.
- [18] S. eka Y. Putri, "Penerapan Model Naive Bayes Untuk Memprediksi Potensi Pendaftaran Siswa Di Smk Taman Siswa Teluk Betung Berbasis Web," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 1, no. 1, pp. 93–99, 2020, doi: 10.33365/jatika.v1i1.228.