I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Menurut Hardle (1994) analisis regresi adalah suatu metode analisis data yang menggambarkan hubungan antara variabel respon Y dengan satu atau beberapa variabel prediktor X. Model regresi umum dari sejumlah n data pengamatan independen $\{(X_i, Y_i)\}_{i=1}^{n}$ adalah:

$$Y_i = m(X_i) + \varepsilon_i \qquad , i = 1, 2, \ldots, n$$

dengan $m(X_i)$ merupakan fungsi regresi atau kurva regresi, ε_i adalah galat yang diasumsikan independen dengan mean 0 dan variansi σ^2 (konstan).

Jika fungsi $m(X_i)$ diketahui distribusinya maka untuk mengestimasi $m(X_i)$ digunakan pendekatan parametrik. Sebaliknya, apabila fungsi $m(X_i)$ tidak diketahui distribusinya maka untuk mengestimasi $m(X_i)$ digunakan pendekatan nonparametrik (Hardle, 1994). Untuk pendekatan nonparametrik, metode yang dilakukan yaitu dengan menggunakan teknik pemulus terhadap data pengamatan. Teknik pemulus yang sudah populer diterapkan seperti histogram, deret fourier, deret orthogonal, dan estimator kernel, spline, serta pengembangan dari fourier yaitu wavelet.

Suatu ukuran kebaikan penduga $m(X_i)$ dapat dilihat dari tingkat kesalahannya. Semakin kecil tingkat kesalahannya semakin baik dugaannya. Salah satu ukuran kesalahan dapat dilihat dari nilai *Mean Square Eror* (MSE). Selain nilai MSE yang kecil peneliti juga mengharapkan mendapat kurva dugaan yang mulus. Namun, teknik pemulus terkadang menghasilkan kurva dugaan yang terlalu mulus atau terlalu kasar sehingga tidak sesuai dengan kurva yang akan diduga. Oleh karena itu, diperlukan teknik pemilihan *bandwidth* (parameter pemulus) optimal yang dapat menghasilkan kurva dugaan yang mulus namun dengan kesalahan minimum.

Terdapat beberapa kriteria pemilihan *bandwidth* yang telah diperkenalkan, seperti *Cross-Validation* (CV), *Generalized Cross Validation* (GCV), *Bayesian Information Criterion* (BIC), *Minimum Description Length* (MDL), *Akaike Information Criterion* (AIC) serta *Improved Akaike Information Criterion* (AIC_C).

Suparti dan Sudargo (2005) telah meneliti kebaikan estimator kernel dan deret fourier dengan mengkaji laju *Integral Mean Square Error* (IMSE). Alifia (2008) telah menerapkan metode Nadaraya-Watson dalam penentuan kurva regresi nonparametrik, sedangkan Sukarsa dan Srinadi (2012) telah meneliti tentang estimator kernel dengan fungsi kernel yang berbeda. Kemudian Rudianto (2015) melakukan penelitian dengan membandingkan bentuk kurva regresi nonparametrik menggunakan metode Fourier dan Nadaraya-Watson pada fungsi regresi linear dan nonlinear dengan pemilihan *bandwidth* menggunakan *Generalized Cross Validation* (GCV).

Dalam penelitian ini, akan digunakan metode Fourier dan Nadaraya-Watson untuk menduga kurva regresi linear dan nonlinear dengan pemilihan *bandwidth* optimal menggunakan AIC, $AIC_{C,}$ dan GCV, serta menentukan metode pemilihan *bandwidth* AIC, AIC_C atau GCV yang paling baik digunakan pada metode Fourier dan Nadaraya-Watson.

1.2 Tujuan

Tujuan dalam penelitian ini adalah:

1. Menduga kurva regresi linear dan nonlinear menggunakan metode Fourier dan Nadaraya-Watson dengan pemilihan *bandwidth* optimal menggunakan AIC, $AIC_{C,}$ dan GCV.
2. Menentukan metode pemilihan *bandwidth* optimal AIC, AIC_C atau GCV yang paling baik digunakan pada metode Fourier dan Nadaraya-Watson dalam menduga kurva regresi linear dan nonlinear.

1.3 Manfaat

Manfaat dari penelitian ini adalah:

1. Menambah refrensi mengenai pemilihan *bandwidth* optimal menggunakan AIC, AIC_C, dan GCV dalam menduga kurva regresi linear dan nonlinear dengan metode Fourier dan Nadaraya-Watson.
2. Menambah pengetahuan tentang metode pemilihan *bandwidth* AIC, AIC_C atau GCV yang paling baik digunakan pada metode Fourier dan Nadaraya-Watson dalam menduga kurva regresi linear dan nonlinear.