

BAB II

TINJAUAN PUSTAKA

2.1. Landasan Teori

2.1.1. Ketergantungan Spasial (*Spatial Dependence*)

Ketergantungan spasial merupakan fenomena di mana nilai suatu variabel pada suatu lokasi tidak sepenuhnya independen dari nilai variabel tersebut di lokasi lain yang berdekatan. Hal ini menjadi ciri khas data geografis, di mana berbagai faktor seperti karakteristik sosial, ekonomi, atau lingkungan di suatu wilayah sering kali berkorelasi dengan wilayah di sekitarnya. Sebagai contoh, tingkat pengangguran di suatu kabupaten dapat dipengaruhi oleh kondisi pasar tenaga kerja di kabupaten tetangga, atau kualitas lingkungan di suatu lokasi dapat mencerminkan kondisi wilayah-wilayah yang berbatasan dengannya. Ketergantungan spasial ini menjadi penting untuk diperhitungkan dalam analisis statistik karena jika diabaikan, estimasi model regresi dapat menghasilkan hasil yang bias dan kurang akurat. Hal ini terjadi karena asumsi independensi residual yang mendasari banyak metode regresi konvensional akan dilanggar, sehingga memengaruhi validitas hasil (Anselin, 2021; Yu et al., 2020). Oleh karena itu, pendekatan yang mampu menangkap ketergantungan spasial menjadi krusial dalam analisis data geografis.

Salah satu metode yang paling umum digunakan untuk mengukur ketergantungan spasial adalah *Moran's I*, yang mengevaluasi apakah ada pola spasial signifikan dalam data. Nilai *Moran's I* berkisar dari -1 hingga 1, dengan nilai positif menunjukkan adanya ketergantungan spasial positif, di mana lokasi dengan nilai serupa cenderung berdekatan, sedangkan nilai negatif mengindikasikan ketergantungan spasial negatif, di mana lokasi dengan nilai berbeda cenderung saling berdekatan (Islam et al., 2022). Proses identifikasi pola spasial juga dapat dilakukan melalui pendekatan pemfilteran *eigenvector*, seperti yang digunakan dalam ESF. Metode ini memisahkan komponen spasial dari residual dengan menggunakan vektor

eigen yang diperoleh dari matriks bobot spasial. Teknik ini memungkinkan analisis yang lebih dalam terhadap pola spasial dengan mengurangi autokorelasi spasial dalam residual, sehingga menghasilkan model yang lebih akurat (Murakami et al., 2024; Murakami & Griffith, 2023). Integrasi pendekatan seperti *Moran's I* dan ESF tidak hanya membantu mendeteksi tetapi juga mengeliminasi efek ketergantungan spasial, menjadikannya alat penting dalam analisis data geografis yang kompleks.

2.1.2. Model Regresi Spasial dan Metode Utama

Regresi spasial merupakan teknik statistik yang dirancang untuk memperhitungkan pengaruh lokasi atau kedekatan geografis dalam analisis hubungan antara variabel dependen dan independen. Data geografis dalam beberapa kasus tidak memenuhi asumsi independensi antar observasi karena adanya keterkaitan spasial. Hal ini dapat disebabkan oleh efek tetangga, di mana karakteristik suatu wilayah dipengaruhi oleh wilayah-wilayah yang berdekatan, atau karena kesalahan spasial yang terstruktur dalam data. Regresi spasial menjadi penting karena model regresi linier konvensional sering kali tidak cukup untuk menangkap pola keterkaitan ini, sehingga menghasilkan estimasi yang bias atau tidak efisien (Dube & Legros, 2014). Regresi spasial dengan mempertimbangkan efek lokasi memungkinkan analisis yang lebih mendalam terhadap pola hubungan variabel yang dipengaruhi oleh dimensi spasial.

Beberapa metode utama dalam regresi spasial telah dikembangkan untuk mengatasi masalah autokorelasi spasial dan heterogenitas lokal dalam data. Salah satunya adalah *Spatial Lag Model (SLM)*, yang memasukkan pengaruh tetangga dalam variabel dependen melalui lag spasial. Metode ini digunakan ketika pengaruh variabel dependen di suatu lokasi dipengaruhi secara langsung oleh variabel dependen di lokasi-lokasi sekitar. Selain itu, *Spatial Error Model (SEM)* digunakan untuk menangkap kesalahan spasial yang terstruktur dalam data, di mana pola spasial tercermin dalam residual model. Pendekatan lain, seperti *Eigenvector Spatial Filtering (ESF)*, menggunakan dekomposisi matriks bobot spasial

untuk mengeliminasi autokorelasi spasial dalam residual, sehingga menghasilkan model yang lebih valid secara statistik. Terakhir, *Spatially Varying Coefficients (SVC)* memungkinkan koefisien regresi bervariasi di setiap lokasi geografis, menangkap heterogenitas lokal yang mungkin ada dalam hubungan antara variabel independen dan dependen.

Regresi spasial tidak hanya memberikan hasil estimasi yang lebih akurat tetapi juga membuka wawasan baru dalam analisis data spasial. Model ini sangat relevan dalam konteks penelitian sosial, ekonomi, dan lingkungan, di mana dimensi geografis sering kali menjadi faktor penting dalam memahami pola hubungan antar variabel. Regresi spasial dalam penelitian ini menjadi alat utama untuk mengeksplorasi dan memahami distribusi Tingkat Pengangguran Terbuka (TPT) di Jawa Barat, dengan memanfaatkan pendekatan seperti ESF dan SVC untuk menangkap pola keterkaitan spasial yang kompleks dan heterogenitas lokal.

2.1.2.1. *Moran Eigenvector*

Moran Eigenvector adalah alat statistik untuk mendeteksi autokorelasi spasial (Griffith & Chun, 2019). Vektor eigen digunakan untuk menangkap pola spasial dominan (Paula-Souza & Diniz-Filho, 2020). Persamaan untuk menghitung *Moran's I*:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n C_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum_{i=1}^n \sum_{j=1}^n C_{ij} \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (2.1)$$

dalam persamaan ini, n adalah jumlah unit spasial, C_{ij} adalah elemen matriks bobot spasial terpusat, Y_i dan Y_j adalah nilai variabel pada lokasi i dan j , serta $\bar{Y}$ adalah rata-rata Y . Indeks *Moran's I* memberikan gambaran tentang tingkat autokorelasi spasial dalam data. Jika nilai $I > 0$, itu menunjukkan autokorelasi positif, di mana wilayah yang berdekatan memiliki nilai yang serupa. Sebaliknya, $I < 0$ menunjukkan autokorelasi negatif, di mana wilayah yang berdekatan memiliki nilai yang berbeda. Nilai $I = 0$ mengindikasikan tidak ada autokorelasi spasial. Vektor eigen yang dihasilkan dari dekomposisi matriks bobot spasial digunakan untuk

memodelkan pola spasial yang dominan dalam data, mendukung analisis spasial seperti ESF dan SVC (Fang et al., 2019).

2.1.2.2. *Eigenvector Spatial Filtering (ESF)*

Eigenvector Spatial Filtering (ESF) adalah metode yang digunakan untuk mengatasi ketergantungan spasial dalam model regresi. ESF mengidentifikasi komponen spasial dengan menggunakan *eigenvector* dari matriks bobot spasial, yang mewakili kedekatan antar wilayah (Abate, 2017; Yoshida et al., 2024). *Eigenvector* yang dihasilkan menggambarkan pola spasial independen dan memungkinkan model untuk mengeliminasi ketergantungan spasial pada residual. ESF adalah metode deterministik berbasis *Moran Coefficient (MC)* dan dapat menguraikan variasi dalam data menjadi komponen yang terkait dengan struktur spasial dan non-spasial (Griffith et al., 2022; Ord, 2024). Persamaan utama model ESF dapat dituliskan sebagai:

$$y = X\beta + E\gamma + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \quad (2.2)$$

y adalah variabel dependen, X adalah matriks variabel independen, dan β adalah koefisien regresi untuk variabel tersebut. Matriks E adalah koleksi vektor eigen yang dihasilkan dari dekomposisi matriks bobot spasial (W), sedangkan γ adalah koefisien yang sesuai dengan vektor eigen E . Komponen residual (ε) mewakili kesalahan yang tidak dijelaskan oleh model.

Guna menghasilkan vektor eigen (E), matriks bobot spasial W pertama-tama diubah menjadi bentuk yang terpusat dengan persamaan $C = \left(I - \frac{11^T}{n}\right) W \left(I - \frac{11^T}{n}\right)$, di mana I adalah matriks identitas, dan 1 adalah vektor satuan. Selanjutnya, matriks C didekomposisi menjadi vektor eigen (E) dan nilai eigen (λ) menggunakan persamaan $C = E\Lambda E^T$. Nilai eigen yang signifikan dipilih berdasarkan kriteria tertentu, seperti rasio $\frac{\lambda_i}{\lambda_j} > 0,25$, untuk memastikan bahwa hanya pola spasial yang dominan yang dimasukkan dalam model. ESF unggul dalam mengurangi autokorelasi

spasial pada residual, menangkap pola spasial yang kompleks tanpa memerlukan asumsi *bandwidth*, seperti pada *Geographically Weighted Regression* (GWR) (Liu et al., 2022).

2.1.2.3. *Random Effects Eigenvector Spatial Filtering* (RE-ESF)

Pengembangan dari model ini yaitu *Random Effects Eigenvector Spatial Filtering* (RE-ESF) mengasumsikan bahwa koefisien untuk eigenvector, yakni γ , bukan lagi ditetapkan secara langsung tetapi diperlakukan sebagai efek acak. Persamaan RE-ESF dituliskan:

$$y = X\beta + E\gamma + \varepsilon, \quad \gamma \sim N(0, \sigma_\gamma^2 \Lambda(\alpha)), \quad \varepsilon \sim N(0, \sigma^2 I) \quad (2.4)$$

vektor γ mengikuti distribusi normal dengan kovarian yang dibentuk berdasarkan *eigenvalue* dari E , ditandai oleh $\Lambda(\alpha)$. Parameter α bertindak sebagai pengendali (*shrinkage* parameter) terhadap kontribusi komponen spasial. Model ini fleksibel karena memungkinkan kontribusi dari masing-masing *eigenvector* menyesuaikan dengan struktur spasial data. RE-ESF memodelkan efek spasial sebagai acak sehingga mampu mencegah *overfitting*, terutama jika jumlah *eigenvector* yang digunakan cukup besar (Murakami, 2017c).

2.1.2.4. *Non-Spatially Varying Coefficients* (NVC)

Model NVC digunakan untuk menangkap dinamika perubahan pengaruh variabel bebas terhadap variabel respon, yang tidak bersifat tetap secara global namun juga tidak bergantung pada lokasi geografis. Model ini dinyatakan dalam bentuk:

$$y_i = \sum_{k=1}^K x_{i,k} \beta_{i,k} + f_{MC}(s_i) + \varepsilon_i, \quad \beta_{i,k} = b_k + f(x_i, k) \quad (2.5)$$

koefisien regresi $\beta_{i,k}$ adalah fungsi dari nilai variabel independennya sendiri x_i, k . Fungsi $f(x_i, k)$ menggambarkan adanya variasi lokal terhadap setiap nilai atribut, namun tidak dikaitkan secara langsung dengan lokasi spasial s_i . Ini berarti bahwa model tetap mempertahankan sifat non-spasial, tetapi memberikan keleluasaan terhadap variasi pengaruh antar nilai variabel

bebas (Murakami, 2017b). Model ini efektif ketika terdapat hubungan non-linear yang tidak bisa ditangkap oleh koefisien regresi tetap.

2.1.2.5. Model *Spatially Varying Coefficients* (SVC)

Model SVC memungkinkan koefisien regresi untuk berubah sesuai lokasi, menangkap variasi spasial yang lebih rinci dalam data. SVC mengasumsikan bahwa pengaruh variabel independen tidak seragam di seluruh wilayah, yang bermanfaat untuk memahami perbedaan efek variabel di berbagai lokasi (Hu et al., 2020; Murakami et al., 2017). Persamaan dasarnya dalam model SVC adalah:

$$y_i = \sum_{k=1}^K x_{i,k} \beta_{i,k} + f_{MC}(s_i) + \varepsilon_i, \quad \beta_{i,k} = b_k + f_{MC,k}(s_i) \quad (2.6)$$

$x_{i,k}$ adalah nilai variabel independen ke- k pada lokasi i , dan $\beta_{i,k}$ adalah koefisien regresi yang bervariasi berdasarkan lokasi s_i . Variasi lokal ini dimodelkan sebagai $\beta_{i,k} = b_k + f_{MC,k}(s_i)$, di mana $\beta_{i,k}$ adalah koefisien global yang berlaku untuk semua lokasi, dan $f_{MC,k}(s_i)$ adalah fungsi spasial yang mencerminkan variasi lokal berdasarkan lokasi s_i .

Keunggulan utama SVC adalah kemampuannya untuk menangkap pola heterogenitas lokal yang tidak dapat dijelaskan oleh model global. Fungsi $f_{MC,k}(s_i)$ biasanya diestimasi menggunakan basis spasial seperti *eigenvector* yang berasal dari matriks bobot spasial. Kombinasi ini membuat model SVC mampu menangkap variasi lokal dan pola spasial dengan lebih baik, menjadikannya lebih stabil dibandingkan metode lain seperti GWR. Penambahan komponen $f_{MC}(s_i)$ pada model SVC juga dapat mengeliminasi autokorelasi spasial dalam residual, yang merupakan masalah umum dalam data geografis (LeSage & Pace, 2018).

2.1.2.6. *Spatially and Non-Spatially Varying Coefficients* (SNVC)

Model SNVC merupakan bentuk paling kompleks dan fleksibel dari semua model yang digunakan. SNVC menggabungkan variasi spasial dan non-spasial pada koefisien regresinya, dengan persamaan sebagai berikut:

$$y_i = \sum_{k=1}^K x_{i,k} \beta_{i,k} + f_{MC}(s_i) + \varepsilon_i, \quad \beta_{i,k} = b_k + f_{MC,k}(s_i) + f(x_i, k) \quad (2.7)$$

koefisien regresi $\beta_{i,k}$ dipengaruhi oleh dua hal sekaligus, yaitu variasi lokasi spasial melalui $f_{MC,k}(s_i)$ dan variasi atribut melalui $f(x_i, k)$. Fungsi $f_{MC,k}(s_i)$ merepresentasikan variasi geografis dari pengaruh variabel bebas terhadap TPT, sementara $f(x_i, k)$ menangkap dinamika non-linear berbasis nilai atribut (Murakami, 2017b). Model SNVC menawarkan pendekatan paling menyeluruh untuk menangkap heterogenitas spasial maupun struktural pada data panel geografis seperti TPT kabupaten/kota.

2.1.3. Penggunaan Paket *spmoran* di RStudio dalam Pemodelan

Spasial

Paket *spmoran* di RStudio menyediakan berbagai fungsi untuk estimasi model regresi spasial baik Gaussian maupun non-Gaussian, serta model dengan koefisien yang bervariasi secara spasial. *Spmoran* di RStudio memanfaatkan *eigenvector* dari matriks bobot spasial untuk mengidentifikasi komponen spasial yang dominan. Studi TPT di tingkat kabupaten/kota di Jawa Barat dengan paket ini dapat membantu menerapkan ESF dan SVC, khususnya untuk mengatasi pola spasial yang berpotensi kuat dalam data (Murakami, 2017b).

2.2. Hasil Penelitian Terdahulu

Hasil penelitian yang terdahulu dicantumkan dengan berfokus pada metode penelitian tersebut yang telah dilakukan sebelumnya. Penelitian relevan sebelumnya berguna sebagai pembanding sekaligus bahan referensi bagi penelitian yang sedang dilakukan. Adapun penelitian-penelitian terdahulu yang relevan ialah sebagai berikut:

Tabel 2. 1. Hasil Penelitian Terdahulu

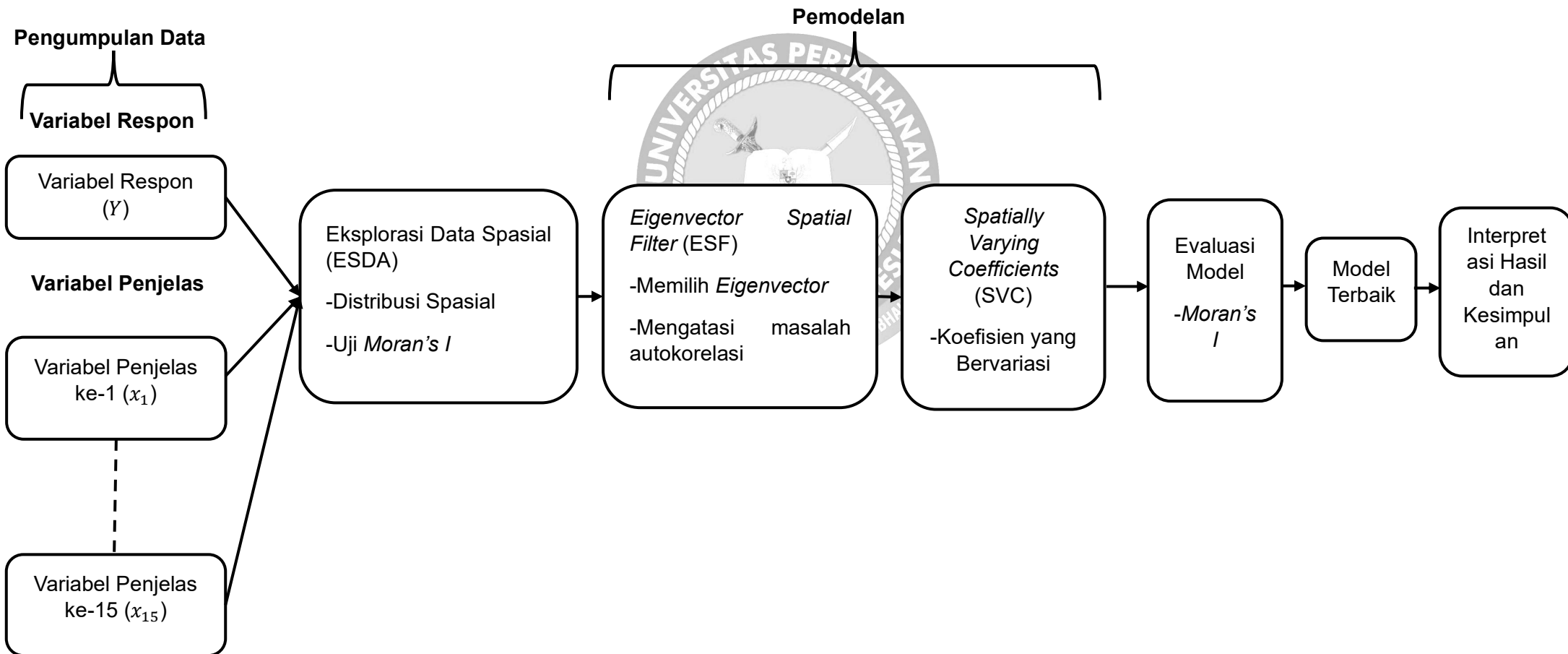
No	Penulis (Tahun)	Judul Penelitian	Metode	Hasil
1	Murakami & Griffith (2015)	<i>Random effects specifications in eigenvector spatial filtering: A simulation study</i>	ESF, Random Effects	Penelitian ini menunjukkan bahwa penggunaan ESF dengan spesifikasi efek acak dapat menangani ketergantungan spasial dalam data besar secara efektif. Hasil simulasi menunjukkan bahwa model ESF dengan efek acak menghasilkan estimasi yang lebih baik dan lebih stabil dibandingkan metode tradisional dalam analisis data spasial yang besar dan kompleks.
2	Murakami (2017)	<i>Spatial regression using the spmoran package: Boston housing price data examples</i>	ESF, SVC	Penelitian ini menjelaskan penerapan ESF dan SVC menggunakan paket spmoran di RStudio. Penelitian ini menggunakan data harga rumah di Boston, hasilnya menunjukkan bahwa ESF mampu mengurangi ketergantungan spasial pada residual, sementara SVC memungkinkan identifikasi variasi pengaruh variabel bebas di berbagai wilayah. Model ini memberikan estimasi yang lebih akurat pada data spasial dengan ketergantungan yang tinggi.

No	Penulis (Tahun)	Judul Penelitian	Metode	Hasil
3	Fotheringham (2017a)	<i>Multiscale geographically weighted regression (MGWR)</i>	MGWR, SVC	Studi ini memperkenalkan MGWR sebagai perkembangan dari GWR untuk menangkap variasi spasial pada berbagai skala. Hasil penelitian menunjukkan bahwa MGWR dapat mengakomodasi koefisien yang bervariasi pada skala lokal dan regional, memberikan fleksibilitas lebih dalam model spasial dan meningkatkan akurasi estimasi pada data dengan heterogenitas spasial.
4	Chen (2022)	<i>Using an Eigenvector Spatial Filtering-Based Spatially Varying Coefficient Model to Analyze the Spatial Heterogeneity of COVID-19 and Its Influencing Factors in Mainland China</i>	ESF-SVC	Penelitian ini menggunakan pendekatan ESF-SVC untuk menganalisis heterogenitas spasial faktor risiko kesehatan terhadap tingkat infeksi COVID-19 di Tiongkok. Model ini menunjukkan performa yang lebih baik dibandingkan OLS, ESF, dan GWR berdasarkan adjusted R^2 yang lebih tinggi dan RMSE yang lebih rendah. Model ini juga mendeteksi perbedaan dampak variabel antar wilayah seperti kepadatan penduduk dan suhu.

Sumber: (Chen et al., 2022; Fotheringham et al., 2017b; Murakami, 2017a; Murakami & Griffith, 2015) yang diolah oleh peneliti

2.3. Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini dapat dilihat pada Bagan 2.1



Bagan 2. 1. Kerangka Pemikiran

Sumber: diolah oleh peneliti