

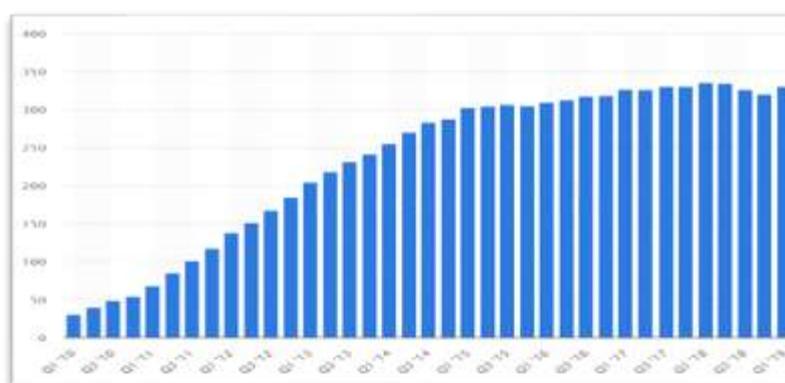
BAB III

METODE PENELITIAN

Bab ini akan menguraikan secara terinci perihal ruang lingkup penelitian, sumber data, metode pengumpulan data, konsep dan definisi yang terkait dengan ruang lingkup penelitian, variabel penelitian, metode analisis penelitian beserta skema analisis yang digunakan. Tahapan tersebut diatas digunakan dalam rangka menjawab seluruh seluruh hipotesis penelitian.

3.1. Ruang Lingkup Penelitian

Penelitian ini bertujuan untuk menganalisis moderasi tingkat kesepakatan investor dan berita atas hubungan antara bias kognitif yaitu efek disposisi dan sentimen investor pada pasar modal di Indonesia. Pengukuran sentimen investor menggunakan platform microblogging Twitter sebagai sumber data penelitian ini untuk *microblog* saham dibandingkan dengan *internet stock message boards* seperti *google finance* atau *Yahoo! finance*. Twitter memiliki penerimaan yang luas di komunitas keuangan dan semua pesan dapat diakses melalui API (*Application Programming Interface*) atau program aplikasi antarmuka situs web. Jika dilihat dari grafik 3.1 disimpulkan bahwa pengguna aktif twitter saat ini dari tahun 2010 semakin meningkat.



Gambar 3.1 Pengguna Aktif Twitter tahun 2010 - 2019

Jumlah pengguna twitter pada kuartal pertama tahun 2019 rata-rata mencapai 330 juta pengguna aktif per bulannya (<https://www.statista.com>). Hal ini menandakan begitu besarnya informasi yang dapat dimanfaatkan oleh investor di pasar modal agar dapat memberikan masukan dalam mengambil keputusan, namun tidak hanya itu saja peluang terjadinya bias psikologis dalam setiap keputusan investasi akan semakin besar meskipun tidak dapat dipungkiri keterbukaan informasi merupakan akses yang dapat dengan mudah didapatkan oleh para investor. Disamping itu terdapat data lainnya yang digunakan dalam penelitian ini yaitu data momentum pada pasar modal serta variabel kebijakan moneter yang dikeluarkan oleh pemerintah Indonesia.

3.2 Objek dan Subjek Penelitian

Objek dalam penelitian ini adalah sentimen investor individu pada pasar modal di Indonesia khususnya investor pengguna platform sosial media yang digunakan untuk mendapatkan informasi ataupun berita perihal situasi yang secara online. Unit analisisnya adalah investor individu pengguna sosial media twitter yang melakukan cuitan, *retweet* dan jumlah pengikut perihal pasar modal di Indonesia. Sedangkan subjek penelitian ini adalah investor individu pengguna sosial media twitter baik yang bertindak sebagai investor, trader maupun pengusaha.

3.3 Desain dan Jenis Penelitian

Dalam penelitian ini digunakan jenis penelitian deskriptif analisis dengan pengujian hipotesis yang menjelaskan fenomena yang diangkat dalam penelitian sehingga menjelaskan model penelitian dengan baik serta mencari kebenaran dan membuktikan hipotesis penelitian dengan menggunakan landasan teoritis untuk membuktikan kebenaran tersebut. Analisis deskriptif juga menggambarkan hubungan yang terjadi antara variabel terikat dengan variabel bebasnya menggunakan metode analisis yang dilakukan secara sistematis sehingga dihasilkan

gambaran yang faktual dan akurat mengenai fakta-fakta atau sifat-sifat serta hubungan antar variabel yang diteliti. Setelah data dikumpulkan kemudian dilakukan pengolahan data yang akhirnya disajikan dalam bentuk tabel atau gambar untuk mempermudah analisis akhir yang dapat menjawab pertanyaan penelitian.

Tujuan dari penelitian ini adalah untuk menemukan dan mengembangkan teori, maka digunakan metode penelitian eksplanasi dengan jenis penelitian eksplanatori survei. Metode ini digunakan untuk membuktikan secara empiris bahwa berita memoderasi hubungan efek disposisi dan sentimen investor yang dimoderasi kembali oleh tingkat kesepakatan investor

3.4 Definisi Operasional Variabel

Definisi operasional variabel adalah penjelasan pada teori sehingga variabel yang digunakan dalam penelitian ini dapat diamati dan diukur untuk mencapai tujuan untuk membuktikan hipotesis awal dalam penelitian ini. Definisi operasional digunakan untuk menjelaskan secara terperinci variabel yang ada dalam penelitian ini untuk menghindari adanya kesalahan penafsiran. Definisi variabel digunakan sebagai landasan berfikir untuk mengungkap atau menjelaskan permasalahan yang diangkat dalam penelitian ini.

Variabel merupakan karakteristik yang dapat diamati dari suatu objek penelitian. Berdasarkan permasalahan diangkat dan hipotesis yang diajukan dalam penelitian ini maka variabel yang digunakan dalam penelitian ini terdiri dari variabel terikat, variabel bebas dan variabel moderasi, dengan rincian sebagai berikut:

1. Variabel terikat atau variabel dependen (Y) dalam penelitian ini adalah sentimen investor.
2. Variabel bebas atau independen (X) adalah efek disposisi.
3. Adapun variabel moderasi primer yang digunakan dalam penelitian ini adalah sebagai berikut:
 - a. *Overreaction* dan *Underreaction*
 - b. Berita Negatif Media

- c. Berita tentang kebijakan moneter
- d. Kualitas Investor Alpha

Dimana keseluruhan variabel moderasi ini diwakili oleh (Wi)

4. Variabel Z yaitu tingkat kesepakatan investor dimana variabel ini memoderasi efek disposisi (X) dan sentimen investor (Y) yang hubungannya dimoderasi oleh variabel moderasi (Wi).

Secara lebih rinci operasional variabel yang digunakan dalam penelitian in adalah sebagai berikut:

Tabel 3.1 Operasionalisasi Variabel

Variabel	Konsep Variabel	Skala/ Pengukuran
Sentimen Investor (Y)	Merupakan sentimen microblogging investor diperoleh dari sinyal atas keputusan investasi berdasarkan sinyal beli, jual atau menahan saham yang dimilikinya. Kemudian sinyal tersebut akan dikonversikan menjadi sinyal “bullish” dan “bearish” yang memperlihatkan kondisi optimis atau pesimis tentang suatu situasi yang dihadapi oleh investor individu dalam data harian. (Sprenger et al., 2014; Antweiler dan Frank, 2004)	Skala/Rasio
Efek Disposisi (X)	Merupakan bias investor dalam keputusan investasi yaitu menjual terlalu cepat saham yang untung dan menahan terlalu lama saham yang rugi (Odean, 1998 ;Goetzmann dan Massa, 2008)	Skala/Rasio
<i>Overreaction/Underreaction</i> (W1)	Merupakan reaksi berlebih atau lesu akibat berita yang disematkan pada harga saham, pendekatan yang dilakukan berdasarkan penelitian yang dilakukan oleh Cohen dan Frazzini (2008)	Skala/Rasio
Berita Negatif (W2)	Berita negatif tentang ekonomi yang di unggah pada platform media maupun otoritas keuangan pada twitter (Mao et al., 2011)	Nominal

Lanjutan Tabel 3.1 Operasionalisasi Variabel

Variabel	Konsep Variabel	Skala/ Pengukuran
Berita Moneter (W3)	Merupakan data berita tentang moneter yang diangkat pada portal berita dan otoritas pada twitter menggunakan tanda tagar #moneter, #sukubunga, #inflasi dengan data harian.	Nominal
Investor Alpha (W4)	Merupakan investor yang memberikan unggahan yang berkualitas dengan pendekatan yang dilakukan oleh Sprenger et al. (2014).	Skala/Rasio
<i>Agreement Level</i> Investor (Z)	Merupakan tingkat kesepakatan antar investor terhadap unggahan pesan yang dituliskan pada twitter (Hong dan Stein, 2007 ; Sprenger et al., 2014)	Skala/Rasio

Adapun penjelasan dari keseluruhan variabel-variabel operasional diatas adalah sebagai berikut:

1. Digunakan sumber data yang berasal dari postingan atau pesan saham pada mikroblog platform twitter.
2. Datanya diakses melalui website menggunakan API (*Application Programming Interface*) yang disediakan oleh twitter.
3. Twitter merupakan *platform social networking services* yang memperbolehkan penggunanya mengirim pesan singkat sebanyak 280 karakter yang disebut tweets dengan jumlah pengguna aktif di Indonesia adalah sebanyak 11,8 juta jiwa.
4. Elemen yang digunakan pada twitter adalah berbentuk hashtag atau “#” yang merupakan kata kunci pesan yang terkandung didalamnya dan dengan sangat mudah untuk diakses.
5. Sama halnya dengan investor yang melakukan postingan pesan atau mencari informasi tentang saham atau bursa saham dengan mudah menggunakan hashtag contohnya adalah “#BBCA”.

6. Data twitter yang diambil berdasarkan cuitan harian yang diunggah pada twitter dengan simbol tagar berdasarkan 70 saham perusahaan yang terdaftar pada bursa saham Indonesia dalam kurun waktu Mei 2015 – Mei 2020. Data tersebut diunggah oleh 558 pengguna twitter yang memberikan opini, informasi maupun berita secara harian. Diperoleh dari 10.533 unggahan berdasarkan tanda tagar (#) kode saham pada 60 perusahaan yang terdaftar pada Bursa Efek Indonesia
7. Data set yang digunakan terdiri dari banyaknya pesan dan akan diklasifikasikan dengan menggunakan metode komputasi linguistik, sejalan dengan penelitian yang dilakukan oleh Antweller dan Frank (2004) dan Sprenger et al. (2014).
8. Untuk metode komputasi linguistik yang digunakan adalah *Naive Bayesian Classification method* yang digunakan secara luas sebagai algoritma dalam mengklasifikasikan teks.
9. Untuk input data menggunakan *Naive Bayesian Classification method* dibuat sampel teks per set nya, dimana setiap teks atau postingan pengguna twitter akan di klasifikasikan sebagai sinyal “buy”, “sell”, “hold”.
10. Setiap klasifikasi akan didefinisikan sebagai bullishness seperti penelitian yang dilakukan oleh Antweller dan Frank (2004), Oh et al. (2011) dan Sprenger et al. (2014) dengan rumus sebagai berikut:

$$Bt = \frac{\ln(1 + Mt^{Buy})}{(1 + Mt^{Sell})}$$

Dimana :

Bt : Bullishness pada periode t

Mt^{Buy} : Pesan periode t untuk “buy”

Mt^{Sell} : Pesan periode t untuk “sell”

Atau Oh et al. (2011) mengadaptasi indeks bullish yang diperkenalkan oleh Antweiler & Frank (2004) yaitu:

$$Bullishness\ Index = \frac{1 + M^{Bull}}{1 + M^{Bear}}$$

Dimana :

M^{Bull} : Total postingan “bullish”

M^{Bear} : Total postingan “bearish”

Indeks ini menunjukkan sentimen investor yang diunggah melalui tweet yang diunggah pada twitter, dimana setiap pesan pada twitter akan dikonversikan sebagai rekomendasi “buy”, “hold”, “sell”. Dimana sinyal “buy”, “hold”, “sell” masing-masing dikonversikan menjadi nilai ‘1’, ‘0’, ‘-1’. Sentimen investor yang diangkat pada penelitian ini adalah sentimen investor microblogging media sosial twitter.

11. Untuk variabel berita media merupakan menggunakan pendekatan yang dilakukan oleh Mao et al., (2011), dengan menghitung jumlah kata dari unggahan berita negatif tentang ekonomi yang diunggah pada platform berita dan otoritas keuangan dan mengambil rasio jumlah kata sentimen negatif dengan jumlah kata dalam tajuk berita. Diunduh melalui portal berita seperti @cnnindonesia, @Bisniscom, @detikfinance, @kompascom dan dua portal otoritas keuangan yaitu @bank_indonesia dan @ojkindonesia, kemudian pada setiap berita yang di posting pada portal diatas digunakan #penurunan, #inflasi dll. Variabel ini merupakan rasio emosional terhadap total jumlah unggahan berita negatif tentang ekonomi sehingga menghasilkan skor sentimen negatif terhadap berita media.

$$Rasio\ Emosional = \frac{Count\ Berita\ Negatif}{Count\ Unggahan\ Berita\ Media}$$

$$NM = \frac{Rasio\ Emosional}{Jumlah\ Artikel}$$

NM : Rasio berita negatif yang diunggah pada twitter

12. Begitupula dengan variabel berita kebijakan moneter pengklasifikasian mengikuti proses sebelumnya pada setiap pesan atau postingan yang berhubungan dengan berita kebijakan moneter dengan menggunakan #moneter, #sukubunga dll. Kemudian untuk periode dengan berita moneter akan di beri label “1” dan periode tanpa berita moneter di beri label “0” hal ini berdasarkan pendekatan yang digunakan Kurov (2010) yang menggunakan data

pengumuman dari FMOC. Variabel ini berdasarkan pada berita kebijakan yang diposting pada twitter melalui portal berita seperti @cnnindonesia, @Bisniscom, @detikfinance, @kompascom dan dua portal otoritas keuangan yaitu @bank_indonesia dan @ojkindonesia

13. Sedangkan investor alpha atau yang secara konsisten memerikan saran yang berkualitas dan mempunyai pengaruh pada sentimen investor, indikasinya adalah *retweet*, *mention* atau *follower* (Sprenger et al., 2014). Untuk mengetahui apakah investor alpha mempunyai pengaruh pada sentimen investor maka setiap tweet atau postingan akan diklasifikasikan sebagai “buy”, “sell”, dan “hold” sebagai sinyal 1, 0, dan -1 (Zhang, 2009). Dengan indeks kualitas sebagai berikut:

$$quality = \begin{cases} = 1 & \text{if } \frac{Sit}{Rit} > 0 \\ = 0, & \text{Otherwise} \end{cases}$$

Sit = Sentimen Investor pada periode t yang berhubungan dengan saham tertentu

Rit = Relative Return pada hari yang sama dengan sentimen investor

Kualitas informasi investor alpha merupakan, informasi yang berkualitas dari investor, trader maupun pengusaha pengguna yang menjadi user aktif pada twitter.

14. *Agreement Level* investor merupakan tingkat kesepakatan antar investor online atau yang disebut sebagai microblogging investor dengan menggunakan pendekatan yang dilakukan oleh Sprenger et al. (2014) yang dinotasikan seperti dibawah ini :

$$At = 1 - \sqrt{1 - \left[\frac{MBuyct - MSellct}{MBuyct + MSellct} \right]}$$

At = Agreement Level periode t

Mbuy = Jumlah Pesan yang dikonverikan kedalam sinyal “buy”

Msell = Jumlah Pesan yang dikonversikan kedalam sinyal “sell”

15. Proksi bobot investor yang terkena dampak efek disposisi berdasarkan

penelitian yang dilakukan oleh Goetzmann dan Massa (2008) yang mengkombinasikan informasi tersebut diatas, sebagai berikut:

$$Wp_{st} = \frac{(S_{Lt} - S_{Gt}) + (B_{Lt} - B_{Gt})}{(S_{Lt} + S_{Gt}) + (B_{Lt} + B_{Gt})}$$

Dimana :

S_{Lt} : *Sell-on-Loss* pada periode t

S_{Gt} : *Sell-on-Gain* pada periode t

B_{Lt} : *Buy-on-Loss* pada periode t

B_{Gt} : *Buy-on-Gain* pada periode t

Namun dalam penelitian ini juga memasukan *hold-on-loss* dan *hold-on-gain* menambahkan perhitungan dari pendekatan Odean (1998), dengan bobot investor yang terkena efek disposisi adalah sebagai berikut :

$$Wp_{st} = \frac{(S_{Lt} - S_{Gt}) + (B_{Lt} - B_{Gt}) + (H_{Lt} - H_{Gt})}{(S_{Lt} + S_{Gt}) + (B_{Lt} + B_{Gt}) + (H_{Lt} + H_{Gt})}$$

Dimana :

H_{Lt} : *Hold-on-Loss* pada periode t

H_{Gt} : *Hold-on-Gain* pada periode t

Sinyal *buy*, *sell*, *hold* yang telah diperoleh di konversikan sesuai dengan transaksi perdagangan *Sell-on-Loss*, *Sell-on-Gain*, *Buy-on-Loss*, *Buy-on-Gain*, *Hold-on-Loss* dan *Hold-on-Gain* pada periode t. Selalu menggunakan harga dimana transaksi investor sebelumnya atau LIFO sesuai dengan periode dan masing-masing kode saham. Semakin besar nilai W_t , maka dampak disposisi pada investor semakin sedikit. Terdapat beberapa periode unggahan tweet yang tidak memiliki Harga awal, maka transaksi perdagangan diambil dari harga periode t-1 atau periode perdagangan aktif sebelumnya.

16. Untuk menghitung bobot *underreaction* dan *overreaction* menggunakan proksi URC sebagai koefisien *underreaction* berdasarkan penelitian yang dilakukan oleh Cohen dan Frazzini (2008), URC merupakan koefisien *underreaction* dengan perhitungan seperti dibawah ini:

$$URC = \frac{RET}{RET + CAR}$$

Dimana :

RET : Return Saham periode t

CAR : Kumulatif Return selama t + 1

Untuk nilai $URC = 1$, maka pasar dalam keadaan efisien

$URC < 1$, menunjukkan adanya respon pasar saham yang lesu atau *underreaction*

$URC > 1$, menunjukkan adanya reaksi berlebihan atau *overreaction*

Data pasar modal yang digunakan adalah harga penutupan per saham serta return per saham dalam periode harian. Data Set Yang Digunakan merupakan data transaksi yang digunakan per hari dengan menggunakan saham-saham yang *listing* pada bursa efek Indonesia atau IDX pada periode periode Januari 2015 – Mei 2020 saham pada 71 perusahaan dengan jumlah microblogging investor sebanyak 558 pengguna.

3.5 Teknik Pengambilan Sampel

3.5.1 Jenis dan Metode Pengumpulan Data

Metode pengumpulan data dilakukan dalam dua tahap yaitu adalah sebagai berikut :

1. Data Sekunder

Jenis data sekunder yang digunakan dalam penelitian ini adalah data transaksi *intraday* menggunakan semua saham yang masing-masing terdaftar pada Indonesia Stock Exchange periode sampel Januari 2015 – Mei 2020 yaitu menggunakan harga dan return saham dimana data ini digunakan untuk mengetahui kualitas investor alpha dan efek disposisi.

2. Data Mining

Data mining bertujuan untuk menarik kata kunci yang diambil dari platform percakapan dalam twitter yang biasa digunakan dalam bentuk “hashtag” atau

hasil yang biasanya merupakan kata kunci yang disertakan dalam banyak pesan untuk dikaitkan atau disebut dengan “tag” dengan topik atau kategori yang akan relevan dan memungkinkan kata kunci ini untuk ditemukan. Data lainnya yang digunakan dalam melakukan analisis dalam penelitian ini adalah kata kunci “Buy” dan “Sell” dan “Hold” atau “bullish”, “bearish”, “netral” yang diambil dari platform percakapan dalam twitter yang biasa digunakan dalam bentuk “hashtag” atau “tag” serta “tweet” atau “retweet” suatu percakapan.

Sedangkan dampak dari berita kebijakan moneter yang dikeluarkan pemerintah data yang digunakan adalah dilihat dari dampak pada saat pasar dalam keadaan *bullish*” dan “*bearish*” menggunakan metode yang dilakukan oleh Kurov (2010).

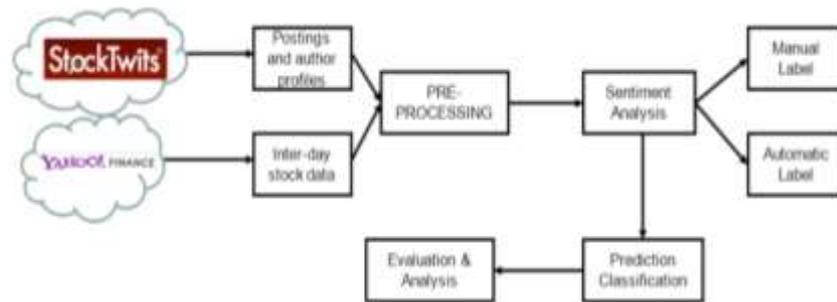
3.5.2 Populasi, Sampel dan Teknik Pengambilan Sampel

Populasi dalam penelitian ini adalah saham yang listing pada bursa saham Indonesia dalam bentuk data harian Januari 2015 – Mei 2020. Pengambilan sampel pada penelitian ini dilakukan dengan menggunakan teknik *purposive sampling* yaitu dengan cara menunjukkan langsung pada suatu populasi berdasarkan karakteristik atau ciri yang dimiliki sampel, dengan tujuan agar diperoleh sampel yang representatif sesuai dengan kriteria yang ditentukan. Kriteria dalam pemilihan sampel untuk data sekunder adalah saham-saham yang menjadi sample dalam penelitian ini adalah harga saham yang masuk kedalam bursa efek indonesia dengan menggunakan data harian dan return saham selama Januari 2015 – Mei 2020.

Sedangkan untuk pengambilan data mining teknik sampling yang digunakan adalah sebagai berikut :

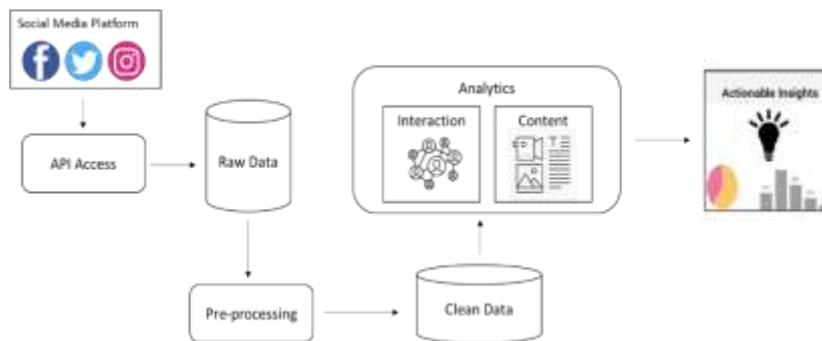
1. Menggunakan teknik yang dilakukan oleh Oh et al. (2011) dengan sistem pipeline yaitu terdapat 5 fase penting yang dilakukan dalam pengambilan data mining yaitu adalah (1) Mengunduh data, (2) Pra pemrosesan, (3) Analisis Sentimen, (4) Klasifikasi prediksi, dan (5) Evaluasi dan analisis.

Dapat dilihat pada gambar pada gambar dibawah ini:



Gambar 3.1 Sistem Pipeline Chong Oh dan Olivia R Liu Sheng (2011)

2. Menggunakan sosial media analytics workflow yang diakses melalui API (Application Programming Interface) atau program aplikasi antarmuka situs web sosial media platform kemudian dilakukan pengunduhan, pra pemrosesan, pembersihan data, di analisis menggunakan *analytics tools* dan akhirnya akan menghasilkan wawasan baru. Dapat dilihat pada gambar pada gambar dibawah ini:



Gambar 3.2 Social Media Analytics Workflow

Sentimen yang diambil adalah dari setiap postingan pada *microblog*, dikarenakan sentimen ini berasal dari *posting* dan tidak secara eksplisit disediakan oleh penulis, harus dilakukan ekstraksi data secara manual dan otomatis.

Dari keseluruhan postingan yang valid hasil dari proses *clean data* maka seluruh postingan tersebut dimasukkan kedalam tiga sentimen berbeda, 1 untuk bullish, -1 untuk bearish, dan 0 untuk sentimen netral yang disebut sebagai "label manual dan digunakan untuk mengevaluasi hasil dari proses pelabelan (sistem) otomatis pada sisa postingan. Seperti halnya (Oh et al., 2011; Kim dan Hovy, 2006;

Schumaker dan Chen, 2009) digunakan pendekatan kedua yaitu label sistem untuk skor setiap kata pada suatu posting dengan leksikon kata kunci *bullish* dan *bearish* yang dibuat secara manual dengan kata *bullish* 1 sedangkan kata *bearish* -1. Skor agregat kemudian diturunkan untuk setiap posting. The bag of words (BOW) digunakan untuk pendekatan diatas.

Untuk memeriksa hubungan antara sinyal dari *microblogs* saham digunakan pengklasifikasian sinyal kunci “Buy” dan “Sell” dan “Hold”. Karena kumpulan data berisi terlalu banyak pesan untuk pengkodean manual, dilakukan pemilihan untuk mengklasifikasikan pesan secara otomatis menggunakan metode mapan dari linguistik komputasi. Sejalan dengan Antweiler dan Frank (2004), menggunakan metode klasifikasi Naïve Bayesian, salah satu algoritma yang paling banyak digunakan untuk klasifikasi teks. Klasifikasi Naïve Bayesian untuk mengkategorikan setiap tweet dan menimbang tweet berdasarkan tingkat kepercayaan berdasarkan jumlah terbobot dari tweet positif dan negatif serta volume tweet (Bartov, et al., 2015).

3.6 Metode Klasifikasi Naïve Bayesian

Metode yang mendasari klasifikasi Naïve Bayesian yang digunakan sesuai dengan pendekatan yang dilakukan oleh Sprenger et al. (2014) dengan notasi sebagai berikut:

$$P(c|d) = \ln P(c) \sum_{1 \leq i \leq nd} \ln P(w_i|c)$$

Dimana :

$P(w_i|c)$: Probabilitas bersyarat dari kata w_i yang muncul dalam dokumen kelas c .

$P(c)$: Probabilitas prior dokumen milik kelas c .

Parameter $P(c)$ dan $P(w_i|c)$ diestimasi berdasarkan set pelatihan dari dokumen yang dikodekan secara manual. Dengan probabilitas sebagai berikut:

$$P(c) = \frac{N_c}{N}$$

Dimana :

N_c : Jumlah dokumen di kelas c dan N adalah jumlah dokumen.

Dengan probabilitas bersyarat $P(w_i|c)$ diperkirakan sebagai berikut :

$$P(w|c) = \frac{W_c}{\sum_{c \in C} W_c}$$

Dimana :

W_c : Jumlah kemunculan kata w dalam dokumen pelatihan kelas c .

Dengan menyertakan *Laplace Smoothing* yang digunakan untuk meminimalkan efek kasus di mana $P(w_i | c) = 0$. Probabilitas bersyarat ini digunakan untuk menggambarkan asumsi bahwa semua kata, atau fitur, tidak bergantung satu sama lain. Kamus dipangkas dengan memilih kumpulan kata yang paling representatif dalam hal kriteria perolehan informasi (IG). IG mengukur perbedaan entropi antara variabel kelas yang tidak dikondisikan dan variabel kelas yang dikondisikan pada ada atau tidaknya kata tersebut. Ini setara dengan informasi timbal balik antara kelas dan kata dan dihitung sebagai

$$IG(w_i, c) = H(c) - H(c|w_i) = \sum_{c \in C} \sum_{w \in \{0,1\}} p(c, w_i) \ln \frac{p(c|w_i)}{p(c)}$$

dimana $p(c, w_i)$ adalah probabilitas gabungan untuk kemunculan kata w_i dan kelas c . Karena penggunaan beberapa kelas, jumlah yang dibobotkan dengan probabilitas masing-masing kelas c dihitung untuk setiap kata.

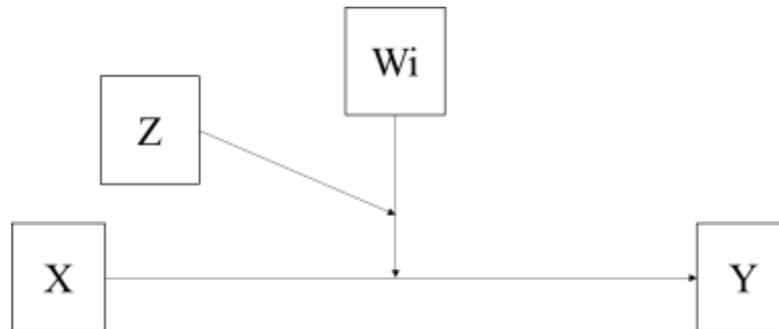
3.7 Pengujian Hipotesis

3.7.1 Variabel dan Analisis Model Regresi Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel terikat, variabel bebas dan variabel moderasi. Variabel terikat atau variabel dependen (Y) dalam penelitian ini adalah Sentimen Investor, sementara itu yang menjadi variabel independen (X) adalah disposition effect. Variabel moderasi terdiri dari overreaction dan underreaction ($W1$), berita negatif ($W2$), kebijakan moneter ($W3$), investor alpha ($W4$) dan untuk variabel moderasi dari efek langsung disposition effect (X) ke sentimen investor (Y) adalah tingkat kesepakatan investor (Z).

Dalam penelitian digunakan uji Regresi berganda dengan analisis moderasi. Pengaruh X dan Y di moderasi oleh W jika ukuran, tanda atau kekuatannya tergantung pada atau dapat diprediksi oleh W. W dikatakan sebagai moderator pengaruh X terhadap Y atau dikatakan bahwa W berinteraksi dalam pengaruh terhadap Y (Hayes, 2018). Mengidentifikasi moderator dari suatu pengaruh membantu untuk menetapkan batasan kondisi dari suatu pengaruh atau keadaan, rangsangan, atau tipe orang yang pengaruhnya besar atau kecil, hadir vs tidak hadir, positif vs tidak ada, positif vs negatif dsb.

Diagram statistik yang sesuai dengan model ini akan membutuhkan bukan hanya dua melainkan tiga atau lebih variabel antesenden dan W serta Z akan menjadi salah satu antesenden tersebut. Untuk itu konseptual diagram pada penelitian ini dapat dilihat pada gambar berikut ini:



Gambar 3.3 Konseptual Diagram

Dimana :

X : Disposition Effect

Y : Sentimen Investor

Wi : *Overreaction* dan *Underreaction*

Berita Negatif

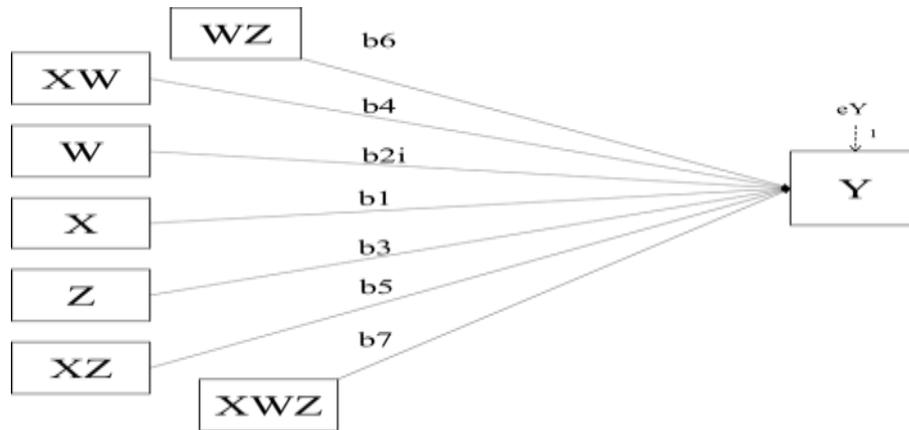
Kebijakan Moneter

Investor Alfa

Z : *Agreement Level*

Sedangkan jika dilihat dari diagram statistiknya dapat dilihat pada gambar berikut

ini:



Gambar 3.4 Diagram Statistik

Untuk itu maka model Regresi berganda bentuknya adalah sebagai berikut:

$$Y = b_1X + b_2W + b_3Z + b_4XW + b_5XZ + b_6WZ + b_7XWZ + e_y \dots \dots \dots 3.1$$

Artinya sebagian besar koefisien regresi diatas mewakili efek kondisionalnya dan tidak boleh ditafsirkan sebagai efek utama, terkecuali untuk koefisien b_7 yang memperkirakan interaksi tiga arah antara variabel X, W dan Z.

Koefisien b_1 , b_2 dan b_3 merupakan efek kondisional atau bisa juga disebut sebagai efek sederhana dan bukan efek utama. Dimana jika dilihat dari persamaan diatas dapat disimpulkan beberapa hal sebagai berikut:

1. Koefisien b_1 , memperkirakan efek X terhadap Y ketika W dan Z bernilai nol.
2. Koefisien b_2 , memperkirakan efek W terhadap Y ketika X dan Z bernilai nol.
3. Koefisien b_3 , memperkirakan efek Z terhadap Y ketika X dan W bernilai nol.
4. Koefisien b_4 memperkirakan interaksi kondisional antara X dan W ketika nilai Z sama dengan nol.
5. Koefisien b_5 , menghitung interaksi kondisional antara X dan Z ketika nilai W sama dengan nol.
6. Koefisien b_6 , memperkirakan interaksi kondisional antara W dan Z ketika X sama dengan nol.

7. Sedangkan rata-rata berpusat pada X, W dan Z.

Terdapat tiga hubungan antar variabel yang jelaskan dalam persamaan diatas yaitu adalah sebagai berikut:

1. Hubungan yang menggambarkan efek kondisional yang diwakili oleh koefisien b1, b2 dan b3. Dimana b1 menunjukkan besarnya pengaruh X terhadap Y ketika W dan Z atau variabel moderasi bernilai nol, b2 menunjukkan besarnya pengaruh W terhadap Y ketika X dan Z bernilai nol sedangkan b3 menunjukkan besarnya pengaruh Z terhadap Y ketika X dan W bernilai nol.
2. Hubungan yang menggambarkan interaksi kondisional yang diwakili oleh koefisien b4, b5 dan b6. Koefisien b4 mengukur pengaruh interaksi kondisional antara X dan W terhadap Y ketika, b5 mengukur pengaruh interaksi kondisional antara X dan Z terhadap Y dan b6 mengukur pengaruh interaksi kondisional antara W dan Z terhadap Y.
3. Hubungan yang menggambarkan interaksi tiga arah antara X, W dan Z terhadap Y, dimana hubungan ini merupakan efek utama dalam menjelaskan pengaruh moderasi W terhadap hubungan X dan Y yang besarnya bergantung pada Z.

Model dalam penelitian ini disebut sebagai model moderasi yang dimoderasi dengan artian dalam hal ini berita dianggap memoderasi hubungan antara disposition effect dan sentimen investor yang dimoderasi oleh tingkat kesepakatan investor pada microblogging investor.

3.7.2 Analisis Data dan Uji Hipotesis

Asumsi yang tidak boleh dilanggar agar penduga tetap BLUE pada metode OLS yang dikenal dengan Teorema Gauss-Markov adalah sebagai berikut :

1. Multikolinieritas

Untuk melakukan pengujian multikolinieritas maka digunakan piranti lunak siap pakai yaitu SPSS *version Process* dengan melihat korelasi antara variabel bebas dengan menggunakan pengujian Correlation. Suatu variabel bebas dikatakan mempunyai hubungan korelasi yang tergolong kuat jika besarnya 0.8

atau lebih. Jika diketahui bahwa variabel bebas yang digunakan mengandung multikolinieritas, maka pada data time series digunakan dilakukan dengan pembedaan atau melakukan differencing pada model regresi dengan menggunakan Ordinary Least Square (OLS)

2. Heteroskedastisitas

Untuk melakukan pengujian terhadap unsur heteroskedastisitas pada model regresi dengan menggunakan Ordinary Least Square (OLS) maka salah satu pengujian yang dapat digunakan adalah residual test dengan menggunakan uji White. Jika nilai perhitungan melebihi nilai kritis α yang dipilih, maka diputuskan bahwa tidak terdapat heteroskedastisitas, sebaliknya jika nilai perhitungan kurang dari nilai kritis α yang dipilih, maka diputuskan bahwa terdapat heteroskedastisitas.

3. Otokolerasi

Dalam melakukan pengujian dimana terdapat unsur otokolerasi, maka dilakukan analisis dengan pengujian Lagrange Multiplier (LM). Jika nilai perhitungan melebihi nilai kritis α yang dipilih, maka diputuskan bahwa model sudah tidak mengandung otokolerasi, sebaliknya jika nilai perhitungan kurang dari nilai kritis α yang dipilih, maka diputuskan bahwa model sudah masih mengandung otokolerasi. Seperti pengujian yang dilakukan pada asumsi-asumsi sebelumnya maka pada pengujian otokolerasi ini akan dengan mudah jika menggunakan piranti lunak siap pakai yaitu *SPSS version Process*.