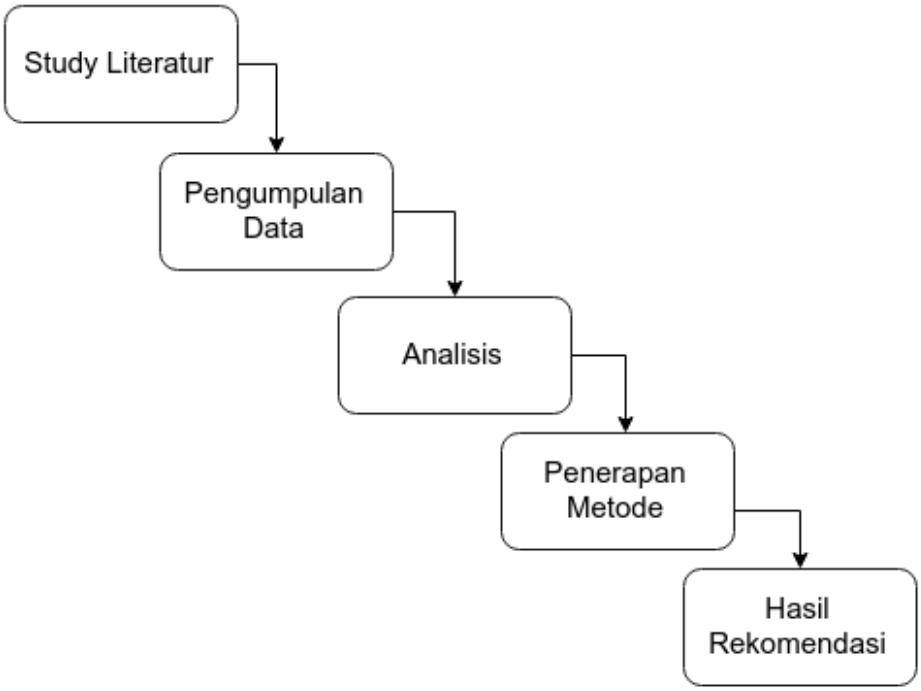
BAB 3

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

Tahapan-tahapan penelitan yang dilakukan dalam melakukan analisis recommender system dengan metode hybrid, antara lain :



Gambar 3.1 Alur Penelitian

1) Study Literatur

Memahami pendekatan yang akan digunakan yaitu metode hybrid, mengumpulkan informasi dari beberapa penelitian yan sudah dilakukan baik dari metode hybrid maupun metode lain yang pernah digunakan.

2) Pengumpulan Data

Mengumpulkan data dari website yang menjadi objek penelitian (yelp.com) data yang diambil seperti : *business, user, reviews, tips,* dan waktu *checkin.*

3) Analisis

Melakukan analisis terhadap sumber data dan pendekatan yang akan digunakan. Menganalisis sejauh mana pendekatan akan berdampak pada sistem rekomendasi yang akan dibuat.

4) Penerapan Metode

Penerapan metode dan pengujian, menghitung nilai kemiripan, nilai prekdisi dan tingkat akurasi dari metode atau pendekatan yang di gunakan pada penelitian ini.

5) Hasil Rekomendasi

Mengemukakan hasil dari uji coba penerapan pendekatan terhadap data yang digunakan. Menampilkan perhitungan seberapa akuratkah *framework* sistem rekomendasi yang di ajukan pada penelitian ini.

3.2 Sumber dan Cara Penentuan Data/Informasi

Dataset Business review di ambil dari website (yelp.com), website tersebut mengadakan kegiatan yang cukup rutin setiap periode yaitu lomba membuat karya ilmiah dari *Dataset* yang bersumber dari sosial media yang dimiliki website tersebut.

Yelp.com merupakan situs sosial media, dimana para pelaku bisnis yang memiliki usah di bidang tertentu seperti makanan, tempat belanja, otomotif dan berbagai bidag usaha lain dapat memasukan informasinya di media sosial tersebut. Selain dapat memasukan informasi mengenai usaha, penggunanya juga dapat melakukan penilaian (*review*) terhadap bidang bisnis yang dimiliki pengguna lain.

3.2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan dataset dari *yelp.com* dimana data merupakan data berkala yang berisi *review* atau penilaian (baik atau buruk) dari para pengguna terhadap suatu bisnis lokal yang terdapat pada sosial media tersebut. Terdapat lebih dari 70 ribu data lokal bisnis dari 10 kota dan 4 negara, untuk data *reviewer* lebih dari 2 juta data yang sudah dikumpulkan dari 2004 sampai dengan 2015.

3.2.2 Cara Penentuan Data dan Informasi

Teknik penentuan data dalam penelitian ini yaitu dengan menentukan populasi, dengan begitu sasarn penelitian menjadi keseluruhan unit. Dari penenttuan populasi maka diambil sampel. Sampel tersebut adalah sebagian jumlah dan karakteristik yang dimiliki oleh populasi atau sebagian populasi untuk dijadikan unit pengamatan sebuah penelitian.

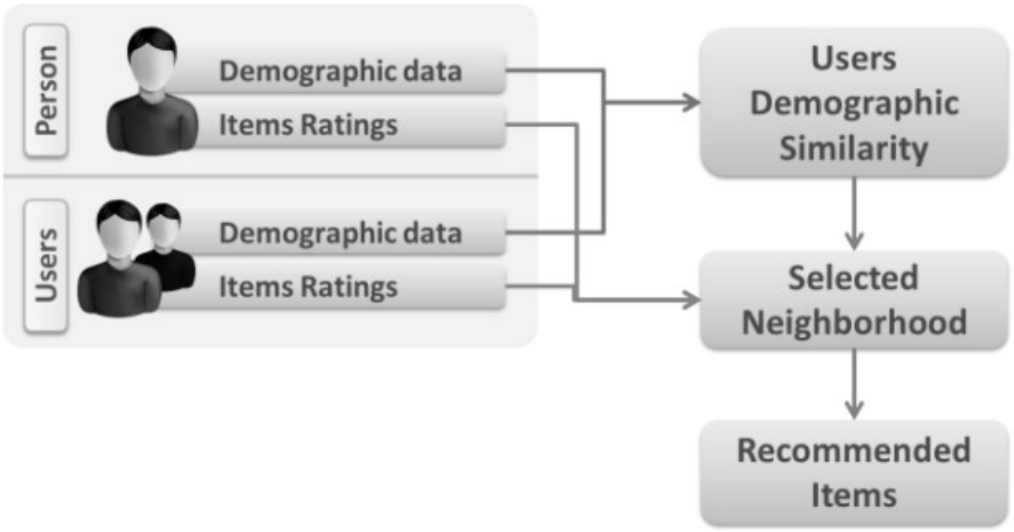
Sesuai pengertian ini maka pengambilan sampel harus diperhatikan agar pemilihan sampel tersebut dapat benar-benar sesuai dengan yang dibutuhkan dalam penelitian dan dapat mewakili populasi.

3.2.3 Perancangan dan Analisis Data

Seperti yang dijelaskan di awal pembahasan bahwa karakteristik user pada penelitian ini dibagi menjadi tiga bagian, dan pendekatan yang akan digunakan pada metode hybrid ini terdapat perbedaan pada ketiganya.

3.2.4 Perancangan Pada User Baru

Pengguna dengan user baru yang belum pernah memiliki aktifitas apapun, akan menggunakan pendekatan yang mersifat *demographic* dan *user ratings,* keduanya akan dikombinasikan untuk kemudian menjadi daftar tempat yang akan direkomendasikan.



Gambar 3.2 *Demographic-based approach*. (Safoury and Salah 2013)

Pada gambar 3.1 diatas menunjukan bahwa sistem rekomendasi yang dibangun untuk User Baru terbentuk dari gabungan demographic user + item ratings. Secara teknis data yang diambil memiliki letak demografi yang berdekatan dan tempat yang memiliki predikat bernilai baik (dinilai oleh pengguna lain).

3.2.5 Perancangan Pada User Dengan Aktifitas Transaksi

Untuk membuat dan menguji pendekatan yang akan digunakan pada user dengan aktifitas transaksi maka dalam penelitian ini sampel yang cocok untuk digunakan yaitu user dengan jumlah transaksi yang sudah banyak dan akan dilakukan pergeseran periode pengambilan data.

3.2.6 Menghitung Nilai Kemiripan

Nilai kemiripan atau Similarity adalah menilai kemiripan user atau item (tempat usaha) sehingga setelah didapatkannya daftar-daftar rekomendasi dari pendekatan yang akan digunakan maka nilai kemiripan ini berperan sebagai nilai urutan dari daftar rekomendasi tersebut. Adapun persamaannya sebagai berikut :

$S(i,j) = \frac{\sum_{u \in U} (Ru,i - \hat{R}i)(Ru,j - \hat{R}j)}{\sqrt{\sum_{u \in U} (Ru,i - \hat{R}i)^2} \sqrt{\sum_{u \in U} (Ru,j - \hat{R}j)^2}}$	Pearson Correlation (3.1)
$S(i,j) = \frac{\sum_{u \in U} (Ru,i - \hat{R}u)(Ru,j - \hat{R}u)}{\sqrt{\sum_{u \in U} (Ru,i - \hat{R}u)^2} \sqrt{\sum_{u \in U} (Ru,j - \hat{R}u)^2}}$	Adjust Cosine (3.2)

Keterangan :

S (i,j) = Nilai kemiripan antara *item* i dan *item* j
u ε U = Himpunan user yang melakukan rating baik *item* i dan *item* j
Ru,i = Rating user u pada item i
Ru,j = Rating user u pada item j
R̂u = Nilai rating rata-rata user u
R̂i = Nilai rating rata-rata *item* i
R̂j = Nilai rating rata-rata *item* j

3.2.7 Menghitung Prediksi Rating

Setelah mendapatkan daftar rekomendasi dan melakukan pengurutan yang bantu oleh nilai kemiripan maka tahapan selanjutnya yaitu menghitung prediksi rating yang didapat dari persamaan berikut :

$$P(a,j) = \frac{\sum_{i \in I} (Ra,i \cdot Si,j)}{\sum_{i \in I} |Si,j|}$$ Weighted Sum (3.3)

Keterangan :

P(a,j) = Prediksi *rating item* j oleh *user* a
i ε I = Himpunan item yang mirip dengan *item* j
Ra,i = *Rating user* a pada *item* I
Si,j = Nilai *similarity* antara *item* i dan *item* j

3.2.8 Menghitung Tingkat Akurasi

Tahapan untuk menghitung atau mengukur akurasi antara nilai yang diprediksi dengan nilai yang sebenarnya, pada tahapan ini persamaan yang digunakan adalah *Mean Absolute Error (MAE)* detail persamaannya adalah sebagai berikut.

$\mathrm{MAE} = \frac{1}{n}\sum_{i=0}^{n} \mid \mathrm{pi} - \mathrm{fi} \mid$	Mean Absolute Error (3.4)

Keterangan :

MAE = Nilai Rata-rata kesalahan hitung
n = Jumlah data yang dihitung
p i = Nilai Prediksi pada data ke-i
f I = Nilai sebenarnya pada data ke-i