**ANALISIS PERBANDINGAN / IMPLEMENTASI SISTEM REKOMENDASI FILM ANI ME MENGGUNAKAN PADA METODE RECURENT NEURAL NETWORK / BIDIRECTIONAL LONG-SHORT TERM MEMORY ATTENTION DAN BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS**

****

**Disusun oleh :**

**MUHAMMAD RIZAL**

**123170036**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK INDUSTRI**

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” YOGYAKARTA**

**2022**

# DAFTAR ISI

[DAFTAR ISI 2](#_Toc16724)

[BAB I Pendahuluan 3](#_Toc17907)

[1.1 Latar Belakang........................... 3](#_Toc32388)

[1.2 Rumusan Masalah................................... 5](#_Toc12083)

[1.3 Batasan Masalah......................... 5](#_Toc17652)

[1.4 Tujuan Penelitian.................... 5](#_Toc10105)

[1.5 Manfaat Penelitian................... 5](#_Toc767)

[1.6 Tahapan Penelitian.................. 5](#_Toc23516)

[1.6.1 Rencana dan Tahapan Penelitian 6](#_Toc28902)

[1.6.2 Metode Pengembangan Sistem 6](#_Toc27926)

[1.6.3 Pengujian Sistem.......... 7](#_Toc14349)

[1.6.4 Pengujian Penelitian............... 7](#_Toc1090)

[1.7 Sistematika Penulisan 7](#_Toc5444)

[BAB II Tinjauan Pustaka 8](#_Toc19085)

[A. Landasan Teori 8](#_Toc22070)

[B. Kerangka Berpikir 8](#_Toc17245)

[C. Penelitian Sebelumnya 8](#_Toc21364)

[BAB III Metodologi Penelitian 8](#_Toc17593)

[A. Metodologi Penelitian 8](#_Toc21433)

[B. Teknik Pengumpulan Data 8](#_Toc26623)

[C. Teknik Analisis Data 8](#_Toc27216)

[D. Analisis Data 8](#_Toc23842)

[BAB IV HASIL DAN PEMBAHASAN 8](#_Toc12751)

[BAB V KESIMPULAN 8](#_Toc7433)

[Daftar Pustaka 8](#_Toc3459)

[Lampiran 8](#_Toc7118)

# BAB I Pendahuluan

## 1.1 Latar Belakang

Berkembangnya variasi atas kebutuhan hidup manusia di masa sekarang seakan terus berkembang seiring dengan pertumbuhan teknologi informasi serta telekomunikasi. Termasuk bentuk hiburan yang merupakan salah satu bentuk kebutuhan yang tak lepas dari kehidupan manusia, salah satu bentuk dari hiburan tersebut adalah film (Billah, M et al., 2021). Film yang merupakan kombinasi dari audio serta visual juga terdiri dari berbagai jenis seperti Movie, TV, Dokumentasi dan sebagainya. Animasi merupakan salah satu bentuk bagaimana film ditampilkan yang merupakan kumpulan dari frame yang digambar menggunakan tangan yang kemudian diolah komputer menjadi animasi, animasi atau anime merupakan salah satu teknologi perfilman yang telah berkembang lama di jepang (Soni, B et al., 2021) Industri anime pada masa sekarang berkembang secara terus menerus dari tahun ke tahun meskipun sempat terjadi penurunan, berdasarkan Anime Report 2020 yang dibuat oleh Asosiasi Animasi Jepang ukuran pasar selama sepuluh tahun terakhir terus berkembang dengan penjualan sebesar 2,51 trilyun yen.

Beberapa penelitian sebelumnya yang pernah meneliti sistem rekomendasi anime yaitu Nuurshadieq & Wibowo., 2020 menerapkan *collaborative filtering* menggunakan LSTM yang bertujuan mengatasi cold-start yang menghasilkan RMSE sebesar 1.4475 yang menunjukkan penelitian tersebut telah lebih baik dibandingkan metode populer seperti SVD dan KNN. Kemudian ada juga penelitian yang dilakukan oleh Soni, B et al., 2021 menerapakan *hybrid recommendation filtering* yang menggunakan algoritma autoencoder dan clustering spectral menghasilkan RMSE sebesar 0.591 dan 0.349. Soni, B., 2021 juga sudah membandingkan dengan beberapa penelitian lainnya sehingga menyimpulkan penelitian yang ia teliti telah menghasilkan hasil yang lebih baik dibandingkan penelitian lainnya.

Penelitian lainnya juga dicoba oleh Billah, M et al., 2021 yang menerapkan sistem rekomendasi anime berbasis *collaborative filtering* menggunakan PCA dan K-Means yang menghasilkan kompleksitas waktu sebesar 2,999602 serta menghasilkan nilai akurasi MMR (Mean Reciprocal Rank) sebesar 0.5619. Pada penelitian selanjutnya yang diterapkan oleh Vie, J. J. et al., 2017 yang menerapkan metode baru yang mereka beri nama BALSE (Blended Alternate Least Squares with Explanation) yang merupakan kombinasi dari beberapa metode, pada penelitian tersebut mereka melakukan ekstrasi fitur pada poster anime dan manga dalam merekomendasikan anime. Komponen dari BALSE adalah Illustration2Vec, ALS (Alternate Least Squares), LASSO (Least Absolute Shrinkage and Selection Operator) serta *Steins gate* yang merupakan metode untuk mengkombinasikan hasil dari dua metode ALS dan LASSO. Dari penelitian tersebut dihasilkan nilai RMSE sebesar 1.4954±0.004 menghasilkan kesimpulan bahwa prediksi BALSE lebih baik dibandingkan ALS.

Penelitian yang berkaitan dengan sistem rekomendasi juga telah diterapkan oleh peneliti lain pada rekomendasi movie menggunakan metode LSTM dan CNN oleh Wentao et al (2020) yang menghasilkan MSE sebesar 0,876 dan MAE 0,751. Penelitian yang dilakukan oleh Haili et al., 2020 juga sama menggunakan metode LSTM-CNN dengan menggunakan dataset movieLens menerapkan personalisasi movie rekomendasi yang menghasilkan MAE sebesar 0,7224 dan MSE 0,691739 dari penelitian tersebut menghasilkan nilai yang lebih baik dibandingkan dengan hanya menggunakan metode CNN. Pada penelitian lainnya yang diteliti oleh Lund, J et al., 2018 meneliti Rekomendasi Movie menggunakan Deep Learning dengan metode autoencoder menghasilkan MAE dan RMSE sebesar 0.15 dan 0.35 selain melakukan menggunakan teknik evaluasi tersebut, penelitian tersebut juga menggunakan evaluasi dengan pengguna langsung dengan jumlah 100 orang partisipan menghasilkan survey sebesar 71.67% partisipan lebih memilih menggunakan hasil penelitian tersebut.

Penelitian lainnya yang menerapkan sistem rekomendasi menggunakan metode BERT telah diteliti oleh Jeong, C et al., 2020 menerapkan *context-aware recommendation* menggunakan metode BERT dan GCN (*Graph Convolutional Network)* yang menunjukkan hasil yang cukup baik dibandingkan dengan metode yang lain dengan menggunakan beberapa evaluasi yaitu MAP sebesar 0.6189, MPR sebesar sebesar 0.6036, Recall@5 sebesar sebesar 0.6736, Recall@10 sebesar sebesar 0.7109 dan Recall@30 sebesar 0.7814. Peneltian selanjutnya yang menerapkan BERT pada *collaborative item-based* penelitian ini membuktikan bahwa metode BERT berhasil melampaui metode LSTM dengan hasil precision@1 sebesar 0.555, precision@10 sebesar 0.079, Recall@10 sebesar 0.791 dan NDCG@10 sebesar 0.669.

Pada penelitian ini juga akan menerapkan metode *attention* pada metode *recurrent neural network* sebelumnya penelitian yang menerapkan metode *attention* seperti pada penelitian Wu, F et al., 2020 yang menggunakan dataset *news recommendation* yang menerapkan metode *attention* yang membuktikan bahwa metode tersebut akan menghasilkan performa yang lebih baik jika dikombinasikan dengan metode *deep learning* lainnya seperti LSTM atau CNN. Salah satu *state-of-art* yang menerapkan *attention* adalah SASRec yang diteliti oleh Kang, W. C., & McAuley, J., 2018.

Sistem Rekomendasi hadir untuk mempermudah dalam pemilihan judul anime yang sesuai. Sistem rekomendasi sendiri umumnya memiliki tiga teknik yaitu *collaborative filtering*, *content-based filtering,* dan *hybrid filtering*. Pada penerapannya teknik *collaborative filtering* menggunakan data user lain dalam menerapkan sistem rekomendasinya yang berakibat pada permasalahan cold-start kemudian untuk *content-based filtering* menggunakan data *meta* dari item itu sendiri yaitu menggunakan kemiripan antar satu konten dengan yang lainnya pada penerapannya teknik ini juga memiliki kelemahan salah satunya adalah membutuhkan rating yang cukup sebelum sistem rekomendasi dapat memberikan hasil rekomendasi akurat yang sesuai preferensi user tersebut, yang kemungkinan jika hanya terdapat rating ketika user tersebut baru terdaftar di dalam sistem maka sistem belum bisa memberikan hasil rekomendasi yang sesuai (Lops, P et al., 2011) berdasarkan dari salah satu kelemahan tersebut hadirlah teknik *hybrid filtering* yang bertujuan untuk mengurangi kelemahan dari kedua teknik tersebut.

Penelitian ini akan menerapkan teknik *hybrid filtering* dengan kombinasi antara *content-based filtering* dan *collaborative* filtering dengan menggunakan hybrid filtering kelemahan pada salah satu tipe sistem rekomendasi akan teratasi. Salah satunya *Cold-start* bisa diatasi dengan penggunaan *content-based filtering* (Wang, H et al., 2020). Dengan menggunakan deep learning model rekomendasi bisa merepresentasikan relasi antar user dan item dengan mempelajari pada *deep-leavel* jaringan struktur non-linear (Wang, W et al., 2020). Pada penelitian ini akan mengunakan metode *Recurrent Neural Network* telah dibuktikan dari beberapa penelitian yang telah dipaparkan sebelumnya bahwa dengan menggunakan metode *recurrent neural network* sistem rekomendasi yang dihasilkan bisa menghasilkan hasil yang dibutuhkan, pada penelitian ini juga akan menggunakan *attention* yang mana *attention* merupakan metode yang mampu mengolah struktur kalimat *complex* (Kang, W. C., & McAuley, J., 2018). Kemudian penelitian ini akan mencoba menerapkan BERT4REC sebagai perbandingannya untuk membuktikan yang mana yang merupakan metode yang lebih baik diantara keduanya. Alasan menggunakan metode BERT adalah metode tersebut merupakan metode state-of-art yang terbaru yang diteliti oleh Fei Sun et al., (2018). Meskipun begitu BERT memiliki kelemahan yaitu ... dan

Data yang akan digunakan bersumber dari dataset yang ada yang berasal dari situs *kaggle* selain itu dataset yang akan digunakan tidak hanya satu tipe dataset. Dengan diterapkan penelitian ini diharapkan dapat menghasilkan sistem rekomendasi yang mampu menghasilkan rekomendasi anime yang lebih baik dari sistem yang telah ada. Sistem yang mampu merekomendasikan sesuai prefensi dari user tersebut serta mampu memperluas pengalaman user dalam menggunakan sistem rekomendasi ini

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan sebelumnya maka didapatkan rumusan masalah sebagai berikut :

1. Metode sistem rekomendasi mana yang terbaik dari metode yang diperbandingkan yaitu BERT dan RNN dengan *attention* pada pemilihan film anime
2. Bagaimana hasil metode sistem rekomendasi yang diterapkan yaitu BERT dan RNN dengan attention

## 1.3 Batasan Masalah

Penelitian ini memiliki batasan dalam penyelesaiannya, sebagai berikut :

1. Menggunakan data penelitian yang bersumber dari *myanimelist* (kaggle)
2. Data penelitian berbahasa inggris.

## 1.4 Tujuan Penelitian

Penelitian ini memiliki tujuan, sebagai berikut :

1. Mengetahui pengaruh metode attention based pada RNN dalam memberikan hasil rekomendasi.

2. Mengetahui tingkat akurasi penerapan metode attention based pada RNN dan BERT terhadap hasil rekomendasi yang diberikan

3. Mengetahui perbandingan akurasi dari RNN dengan *attention* dan BERT pada sistem rekomendasi

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah berusaha untuk memberikan hasil rekomendasi terbaik kepada orang - orang penonton aktif film anime maupun orang - orang yang baru mengenal film anime, selain itu penelitian juga bertujuan untuk meningkatkan eksplorasi metode deep learning pada sistem rekomendasi.

## 1.6 Tahapan Penelitian

Tahapan penelitian yang diterapkan pada penelitian ini adalah sebagai berikut :

### 1.6.1 Rencana dan Tahapan Penelitian

a. Studi Literatur

Pada penelitian ini permasalahan dan penyelesaian yang diselesaikan dihimpun dari berbagai referensi sumber literatur yang relevan dan sesuai dengan penelitian ini.

b. Pengumpulan dan Pengolahan Data

Pada tahapan ini data dikumpulkan berasal dari situs *kaggle* yang kemudian diolah dan dipilah yang mana saja yang akan dijadikan data penelitian.

c. Analisis Sistem

Analisis sistem ini dilakukan untuk menganalisa berbagai keperluan dalam proses perancangan sistem sehingga memudahkan dalam proses selanjutnya.

d. Perancangan Sistem

Pada tahapan ini merupakan penerapan dari analisis sistem sebelumnya yaitu menerapkan pemodelan terhadap sistem yang akan dibuat pada penelitian ini.

e. Implementasi Perangkat Lunak

Tahapan ini merupakan tahapan mengimplementasikan sistem yang sebelumnya sudah dirancang.

f. Pengujian dan Analisis

Pada tahapan ini dilakukan pengujian terhadap sistem yang telah diimplemtasikan sebelumnya, yang kemudian akan dilakukan analisis terhadap hasil pengujian yang telah dilakukan.

g. Kesimpulan dan Saran

Pada tahapan ini penelitian akan diberikan kesimpulan dari hasil pengujian yang telah dilakukan dan kemudian akan disertakan saran yang selanjutnya dapat dikembangkan untuk penelitian selanjutnya sehingga mendapatkan hasil penelitian yang lebih baik.

### 1.6.2 Metode Pengembangan Sistem

Metode pengembangan sistem yang akan digunakan pada penelitian ini adalah *prototyping.* *Prototyping* atau *prototype* digunakan dengan alasan karena dengan menggunakan metode ini kedekatan antara perancang dan pengguna. Dengan proses yang dilakukan secara terstruktur pada setiap tahapan pembuatannya membuat sistem lebih cepat dan lebih hemat dibandingkan metode pengembangan sistem lainnya. Tahapan pada proses *prototyping* (Pressman, 2015) adalah sebagai berikut :

1. *Communication*

Pada tahapan awal pengembangan sistem melakukan komunikasi dan kolaborasi antara pengguna atau pemangku kepentingan dengan maksud memahami tujuan dan kebutuhan sistem sehingga bisa mempermudah dalam menentukan fitur dan fungsi sistem yang akan dibangun.

2. *Planning*

Pada tahapan ini dilakukan perencanaan yang berfungsi untuk mempermudah pengembang saat proses pembuatan sistem. Tahap planning mendeskripsikan tugas teknis, resiko yang mungkin akan terjadi, kebutuhan sumber daya, hasil produk, dan jadwal pengerjaan sistem.

3. *Model*

Pada tahapan ini pengembang membuat model dari sistem yang akan dibuat sehingga pengembang dapat memahami kebutuhan sistem dan desain yang sesuai untuk menunjang kebutuhan tersebut.

4. *Construction*

Pada tahap construction, pengembang memulai pembuatan sistem tahap sebelumnya, selain itu tahap ini juga melakukan pengujian atau testing untuk menemukan kesalahan pada pembuatan sistem.

5. *Deployment*

Pada tahap ini sistem yang telah dibuat dikirimkan kepada pengguna baik semua fitur selesai maupun sebagian untuk mendapatkan evaluasi produk dan memberikan feed back berdasarkan evaluasi.

### 1.6.3 Pengujian Sistem

Pengujian sistem yang akan digunakan pada penelitian ini adalah metode black box testing, yaitu metode yang pengujian sistem yang menekankan fungsionalitas tanpa mengetahui coding dari sistem tersebut. Black box testing bertujuan untuk mengukur kinerja dari sistem yang telah dibangun.

### 1.6.4 Pengujian Penelitian

Pengujian penelitian biasa digunakan pada penelitian rekomendasi sistem adalah root mean squered error (RMSE) dan Mean Average Error (MAP) (Batmaz, Z et al., 2019).

## 1.7 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam menyusun laporan penelitian ini adalah sebagai

berikut:

**Bab I Pendahuluan**

Pada bagian ini membahas tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metodologi penelitian, dan sistematika penulisan.

**Bab II Tinjauan Literatur**

Tinjauan literatur memuat tentang dasar teori yang sudah ada sebagai bahan referensi terkini dan pondasi untuk memperkuat argumentasi dalam penelitian ini sekaligus mendasari pemecahan masalah dalam penelitian ini.

**Bab III Metodologi Penelitian dan Pengembangan Sistem**

Bab ini membahas tentang tahap perancangan kebutuhan, tahap analisis, dan tahap perancangan serta meberikan gambaran garis besar penyusunan program.

**Bab IV Hasil, Pengujian dan Pembahasan**

Pada bab ini akan menyajikan hasil penelitian berisi hasil implementasi dari perancangan yang telah dibuat pada bab sebelumnya dan berisi pengujian terhadap hasil penelitian beserta pembahasannya

**Bab V Kesimpulan dan Saran**

Pada bab ini berisi kesimpulan dari hasil penelitian dan saran yang diajukan untuk pengembangan penelitian selanjutnya.

# BAB II Tinjauan Pustaka

## Anime

Anime merupakan kartun yang aslinya berasal dari jepang, selain itu istilah Anime dalam bahasa inggris adalah *animation* (B. Soni et al., 2021). Anime atau kartun pada dasarnya memiliki definisi yang sama yaitu kumpulan gambar / frame yang disusun sehingga membentuk animasi. Perbedaan yang membuat anime berbeda pada kartun yang sejenis adalah desain dari karakter anime yang memiliki ciri khas tersendiri seperti memiliki mata yang besar. Anime juga memiliki format penanyangan yang cukup bervariasi yaitu TV, Movie, ONA, OVA, dan OAD (Billah, M et al., 2021). Selain itu anime juga memiliki banyak kategori tidak hanya genre yang bervariasi terdapat musim tayang, sumber cerita (*manga*, *original, light novel)*, dan studio.

## Sistem Rekomendasi

Sistem rekomendasi merupakan sistem yang bertujuan untuk menghasilkan suatu *item* kepada user sehingga user mampu meningkatkan pengalaman penggunaan terhadap aplikasi yang menggunakan sistem rekomendasi tersebut, sehingga user bisa membuat keputusan terhadap sesuatu hal yang diinginkanny,. biasanya sistem rekomendasi menghasilkan item secara spesifik seperti merekomendasikan musik ataupun berita (Ricci, F et al., 2011). Sistem rekomendasi juga bisa dilihat sebagai sistem pencarian ranking dimana query masukannya adalah kumpulan dari user dan konteks data informasi yang berhubungan (Cheng, H. T et al., 2016)

Sistem rekomendasi hadir untuk mengurangi derasnya arus informasi yang hadir di internet dengan adanya sistem rekomendasi user mampu memilih serta menentukan arus informasi yang seperti apa yang sesuai kebutuhan user tersebut. Sistem rekomendasi merupakan sistem yang terpersonalisasi yang ditujukan secara khusus kepada setiap user tergantung dari kebutuhan user itu sendiri sehingga tujuan dari sistem rekomendasi adalah meningkatkan pengalaman ditujukan kepada satu user bukan merepresentasikan suatu grup secara keseluruhan (Burke, R et al., 2011)

Dalam pendekatannya sistem rekomendasi biasanya menggunakan dua pendekatan yaitu *collaborative filtering, content based filtering* dan *knowledged based.* Dengan perkembangan informasi saat ini yang lebih kompleks maka hadirlah metode *hybrid* *filtering* yang mengkombinasikan kedua kemampuan dari metode tersebut sehingga bisa saling menutupi kekurangan satu sama lain.

### Content Based Filtering

*Content Based Filtering* merupakan salah satu pendekatan yang terdapat pada sistem rekomendasi yang lebih berfokus terhadap attribut atau *features* yang terdapat pada *item* yang akan direkomendasikan kepada user. Content based filtering merupakan metode yang dimana orang - orang yang menyukai sebuah item dengan beberapa attribut di aktifitas sebelumnya yang dimana di masa depan akan memiliki item yang rekomendasi yang mirip dengan item tersebut (Çano, E., & Morisio, M., 2017).

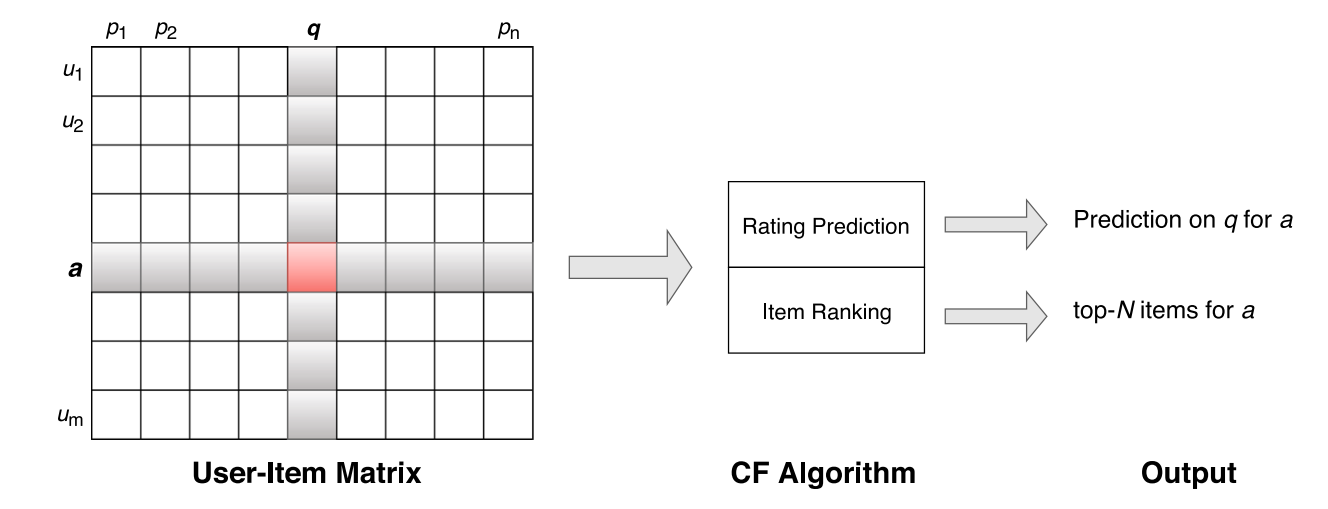
Satu kunci permasalahan dari rekomendasi content-based adalah kualitas dari feature. Item yang direkomendasikan membutuhkan deskripsi attribut yang jelas dan berarti sehingga rekomendasi item tersebut kepada user bisa dilakukan (Burke, R et al., 2011). Teknik *content-based filtering* tidak membutuhkan data *profile* orang lain karena hal tersebut tidak mempengaruhi hasil rekomendasi, selain itu jika data user profle berubah teknik content-based masih mempunyai potensi untuk menyesuaikan hasil rekomendasi dalam waktu yang ingkat (Isinkaye, F et al., 2015).

Keuntungan dengan menggunakan teknik content-based yaitu mampu memberikan rekomendasi item baru meskipun user tidak memberikan rating kepada item - item yang dituju oleh user tersebut, jadi meskipun database tidak menyediakan data user preferensi hal itu tidak mempengaruhi tingkat keakurasian sistem rekomendasi (Isinkaye, F et al., 2015). User tidak perlu untuk berbagi preferensi satu sama lain sehingga memastikan keamanan *privacy* user tersebut. *Content-based filtering* juga memberikan rekomendasi yang *transparency* artinya sistem rekomendasi yang diberikan sangat jelas dan masuk akal karena memiliki keterhubungan dengan item yang user sukai dari aktifitas user itu sendiri berbeda dengan *collaborative filtering* yang mengambil kecocokan item berdasarkan data dari user lain yang artinya *item* rekomendasi yang diberikan belum tentu memiliki kesamaan *features* dengan *item* yang sesuai prefensi user tersebut (Ricci, F et al., 2011).

Kekurangan dari menggunakan pendekatan CBF ini adalah Keterbatasan hasil rekomendasi karena terlalu bergantung kepada features yang terdapat pada item preferensi user tersebut sehingga hasil rekomendasi yang diberikan oleh sistem akan terbatas pada konten dengan tema yang sama yang kadang bisa memberikan hasil rekomendasi yang sama tanpa ada hasil rekomendasi yang baru. (Ricci, F et al., 2011). Pada teknik CBF juga sulit mendapatkan *feedback* dari user karena user terbiasa untuk tidak memberikan rating kepada hasil rekomendasi yang diberikan mengakibatkan sulit untuk mengetahui apakah rekomendasi yang diberikan merupakan hasil yang benar atau tidak. (Modallal, S., 2015)

### Collaborative Filtering

Collaborative Filtering adalah metode sistem rekomendasi yang menghasilkan hasil rekomendasi secara spesifik yang dilakukan dengan cara melakukan pendekatan berdasarkan kemiripan selera dengan pengguna lain. Collaborative filtering adalah sebuah sistem rekomendasi yang mempunyai cara kerja yang cukup berbeda dengan content-based pada teknik ini prediksi sistem rekomendasi bukan dihasilkan dari deskripsi metadata seperti deskripsi yang terdapat pada movie dan musik, pada teknik pendekatan yang digunakan adalah menggunakan preferensi item oleh user jadi user yang memiliki kemiripan preferensi akan menghasilkan prediksi sesuai kemiripan preferensi antar user. (Isinkaye, F et al., 2015) Aktor utama dari sistem rekomendasi CF adalah user yang secara aktif mencari hasil rekomendasi dari ranking items atau prediksi rating. Dengan memanfaatkan prefensi sebelumnya sebagai patokan untuk menentukan korelasi antar user, sebuah pendekatan CF mengandalkan prefensi pengguna yang sesuai. (Batmaz, Z et al., 2019) .



**Gambar 2.1** Gambaran sederhana Colllaborative Filtering (Batmaz, Z et al., 2019)

Umumnya terdapat pada collaborative filtering terdapat dua pendekatan yaitu *memoery-based* dan *model-based.*

1. Memory-based filtering

*Memory based filtering* dikenal sebagai *neighborhood fltering*, metode collaborative filtering memanfaatkan kombinasi dari *ratings user-items* yang terprediksi melalui basis kedekatan mereka (Aggarwal, C. C., 2016). Terdapat dua pendekatan yang digunakan pada *memory-based filtering* ini yaitu :

1. *User-based collaborative filtering*

Pada pendekatan *user-based* ini memanfaatkan kemiripan preferensi user yang kemudian memberikan rekomendasi item berdasarkan kemiripan preferensi user. Sebagai contoh misal user A dan User B menyukai item P maka hasil rekomendasi jika user B menyukai item Q berarti ada kemungkinan user A akan menyukai item Q juga.

1. *Item-based collaborative filtering*

Pada pendekatan *item-based* berfokus pada kemiripan selera item jadi hasil rekomendasi yang diberikan akan berdasarkan kemiripan selera item sebagai contoh misal banyak user yang jika menyukai item A maka menyukai item B maka user yang menyukai item A akan diberikan rekomendasi item B.

1. Model-based filtering

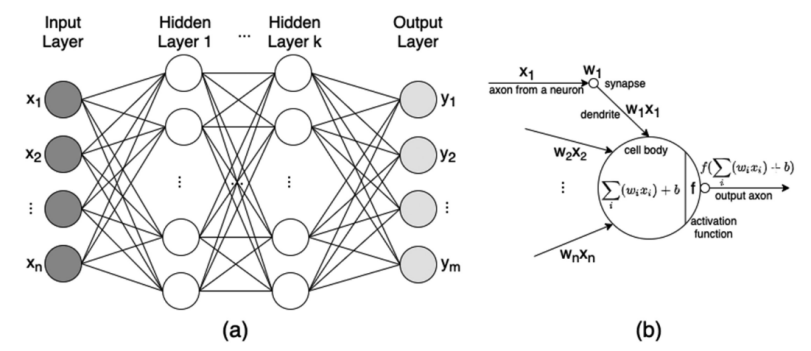
*Model-based filtering* merupakan teknik collaborative filtering yang menggunakan model machine learning atau teknik data mining. Dengan menggunakan pendekatan ini dibandingkan dengan *memory-based* atau *neighborhood-based systems* pada *model-based* pendekatan yang digunakan adalah dengan menggunakan rating untuk belajar pada model prediksi yang akan diterapkan (Ricci, F et al., 2011). Ide dasar dari model-based adalah memodelkan hubungan antara user-item dengan fitur antara pengguna dan item dalam sistem, seperti preferensi dari user dan kategori dari item. Model ini dilatih dengan menggunakan data yang tersedia yang kemudian digunakan untuk memprediksi rating user untuk item baru.

### Hybrid Filtering

Hybrid Filtering yaitu teknik sistem rekomendasi yang menggabungkan atau mengkombinasikan dua atau lebih teknik sistem rekomendasi. Variasi teknik rekomendasi yang telah ada hingga saat ini adalah *content-based, collaborative-filtering, knowledge-based* dan *demographic technique* dari tiap teknik tersebut masing - masing memiliki kekurangan sehingga untuk menanggulangi kekurangan tersebut adalah dengan menggunakan teknik *hybrid filtering*. Salah satu kelemahan yang paling umum dari teknik *content-based* dan *collaborative filtering* adalah *cold-start* (Burke, 2007) yang dimana kelemahan tersebut adalah singkatnya adalah apa hal yang direkomendasikan kepada user yang memiliki hanya sedikit rating.

### Deep Learning

Deep learning merupakan bidang yang masih sangat digandrungi oleh komunitas *machine learning* dan *data mining*, *deep learning* itu sendiri termasuk cabang dari machine learning (Alfarhood, M., & Cheng, J. 2021). Model dari *deep learning* ini sendiri bisa dilatih dengan *supervised learning* ataupun unsupervised learning. Definisi sederhana dari *deep learning* adalah setiap jaringan syaraf dengan lebih dari dua lapisan (Suyanto et al., 2019). Berikut gambaran dasar dari deep learning :



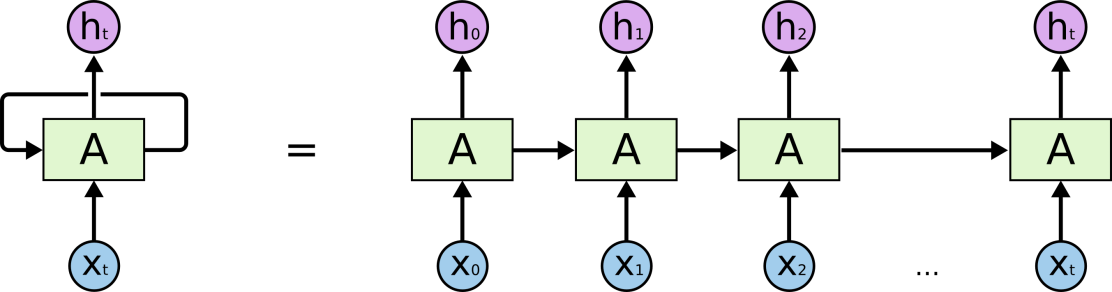
**Gambar 2.2** **(a) Arsitektur umum dari jaringan syaraf deep learning, (b) Jaringan syaraf tiruan: Diagram Perhitungan dasar untuk jaringan syaraf. (Selvaraj, S., 2021)**

Alasan mengapa *deep learning* cocok digunakan pada bidang sistem rekomendasi menurut Alfarhood, M., & Cheng, J. 2021 adalah

1. E*nd to End Differentiable* maksudnya adalah umumnya pada metode tradisional penelitian dilaksanakan pada beberapa tahap dengan menggunakan *deep learning* hal tersebut bisa diperpendek dengan menjadikannya hanya dengan satu tahapan yaitu dengan hanya menggunakan sebuah *neural network.*
2. Memberikan *inductive bias* yang sesuai kepada tipe data masukan. *Inductive bias* memungkinkan sebuah algoritma pembelajaran untuk memprioritaskan satu solusi dibandingkan yang lain, *inductive bias* mampu memberikan asumsi tentang proses pembuatan data atau kemungkinan ruang solusi (Battaglia, P. W et al., 2018). Karena hal tersebut jika terdapat struktur yang inheren yang dapat dieksploitasi oleh model, maka dengan begitu deep neural networks bisa berguna.

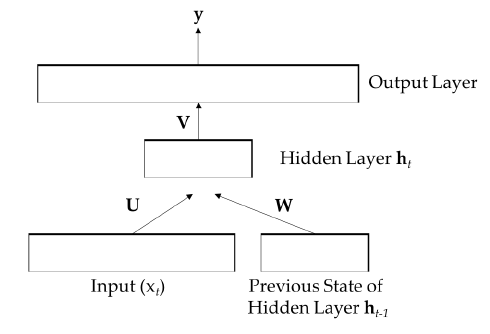
### Recurrent Neural Network

*Recurrent Neural Network* (RNN) merupakan salah satu jenis arsitektur jaringan syaraf tiruan, RNN sendiri merupakan anggota keluarga dari *neural network* untuk menangani data yang berkelanjutan atau bersambung (*sequential data)* (Goodfellow, Ian et al., 2016). Berbeda dengan CNN yang digunakan secara efektif pada pengolahan data *spatial* Dengan menggunakan RNN yang didesain untuk menangani data sequential lebih baik (Gao, Z., & Wang, X., 2019). RNN disebut *reccurent* karena dalam prosesnya RNN melakukan proses yang sama pada setiap elemen dalam urutan, dengan keluaran yang bergantung pada perhitungan sebelumnya. Pada RNN konsepnya adalah bagaimana menangani data yang saling berhubungan satu sama lain atau data yang berututan atau bersambung (*sequential).* RNN memperkenalkan variabel status yang menyimpan informasi masa lalu, bersama dengan *input* saat ini serta untuk menentukan *output* saat ini. RNN merupakan metode yang berfokus pada sifat data dimana instance pada waktu sebelumnya yaitu (t-1) memberi pengaruh pada instance waktu selanjutnya (t), karena hal tersebut RNN mampu mengingat masa lalu (Jan Wira Gotama Putra, et al., 2020).



**Gambar 2.3 diagram recurrent neural network.**

Dalam bentuk matematisnya jika diberikan sebuah sekuens input . Data yang bisa merupakan data vektor, gambar, teks ataupun suara dipengaruhi oleh data sebelumnya yang bisa ditulis sebagai (Jan Wira Gotama Putra, et al., 2020). Meskipun RNN digunakan dengan tujuan untuk mengingat kejadian secara keseluruhan, tetapi secara praktikal hal itu sulit dilakukan dalam urutan kejadian yang panjang, hal ini ini dikenal dengan *vanishing* atau *exploding gradient problem*. Berikut konsep dasar dari *Reccurent Neural Network* :



**Gambar 2.4 Konsep Neural Network (Jan Wira Gotama Putra, et al., 2020)**

Berdasarkan konsep sederhana pada diagram diatas, konsep tersebut sudah sessuai dengan tujuan dari konsep RNN yaitu mengingat kejadian sebelumnya (Jan Wira Gotama Putra, et al., 2020). Konsep RNN di atas bisa ditulis kembali dalam persamaan sebagai berikut :

……………………………………………….(2.1)

Keterangan :

: *hidden state* ke-t

: Fungsi *activation*

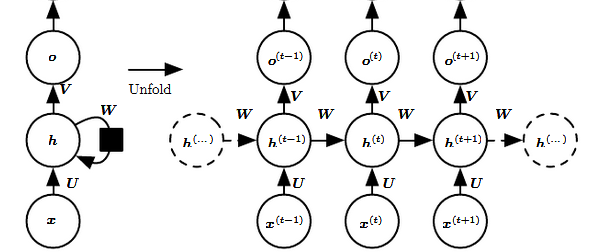
: Nilai *Input* ke-t

: hidden state ke t-1

b : bias

Dari persamaan diatas didapat *f* merupakan fungsi aktivasi yang adalah non-linear dan dapat diturunkan (Jan Wira Gotama Putra, et al., 2020). *b* adalah *bias* yang merupakan parameter yang selalu diikut sertakan kedalam *artficial neural network.* Konsep pada persamaan di atas punya analogi dengan *full markov chain* artinya, *hidden state* pada saat ke-t bergantung pada semua *hidden state* dan input sebelumnya. Fungsi *f* bisa diganti dengan fungsi variasi RNN lainnya misal nya bisa menggunakan *long short-term memory* (LSTM).fungsi diatas jika dijabarkan adalah sebagai berikut :

*,*



**Gambar 2.5 Diagram arsitektur RNN (Goodfellow, Ian et al., 2017)**

Pada diagram gambar 2.5 diatas merupakan detail dari gambar 2.3 dari diagram di atas terlhat proses RNN dengan simbol matematisnya, bagian kiri merupakan arsitektur dasar dari RNN sedangkan bagian kanan merupakan penjabaran dari bagian kiri. Sebagai contoh pada permasalahan kalimat dengan tiga kalimat, maka jaringan RNN yang tergambar seperti pada gambar 2.5 satu layer untuk masing - masing kata. Penjelasan dari diagram di atas adalah sebagai berikut :

**-** **Input**: *x(t)* merupakan *input* yang masuk kedalam jaringan pada langkah ke - t.

**-** **Hidden state:** *h(t)* merepresentasikan sebuah *hidden state* pada langkah ke - t yang bertindak sebagai *memory* pada jaringan. Seperti yang sudah dijelaskan sebelumnya fungsi *f* ini bisa digantikan dengan persamaan *non-linear* seperti *tanh, ReLU.*

**- Weights:** *U* merupakan parameter bobot yang terdapat pada hubungan *input* menuju *hidden state* sedangkan *W* merupakan nilai parameter bobot yang menghubungkan antara *hidden state* dengan *hidden state,* dan *V* merupakan parameter bobot yang menghubungkan *hidden* dengan *output.*

**- Output:** *o(t)* merupakan nilai keluaran dari proses yang telah dilakukan oleh RNN pada langkah ke t. Pada keluarna ini jika dirumuskan menjadi

* + 1. **Forward Propagation**

*Forward propagation* merupakan proses yang dilewati setiap *neural network.* Pada proses *forward propagation* ini *neural network* menerima *input* yang nantinya akan diproses oleh *hidden state*. Proses *forward propagation* pada proses *hidden state* diproses melalui *activation function.* Dengan proses *forward propagation* ini akan didapatkan hasil keluaran. RNN merupakan metode yang mengambil prinsip parameter sharing yang mana neuron yang sama akan diulang-ulang pada saat proses feed forward (Jan Wira Gotama Putra, et al., 2020). Pada setiap *neuoran* di *hidden state* atau *output layer* terdapat dua proses yang dijalani yaitu :

1. **Preactivation :** Pada proses ini dilakukan penjumlahan bobot dari *input.*
2. **Activation :** Berdasarkan proses *preactivation* sebelumnya yang merupakan proses penjumlahan bobot dari input, proses *activation* merupakan proses yang menggunakan fungsi matematika yang mana menambah persamaan non-linear ke jaringan, umumnya terdapat fungsi aktivasi popular yang digunakan yaitu *sigmoid, hyperbolic tangent(tanh), ReLU* dan *Softmax.*

Berikut merupakan formula persaman yang merepresentasikan *forward propagation* pada RNN :

…………………………………………………………(2.3)

…………………………………………………………………….(2.4)

………………………………………………………………………(2.5)

………………………………………………………………..(2.6)

Pada formula diatas ***b*** dan ***c*** merupakan vector bias dengan bobot matrik yaitu ***U, V*** dan ***W*** secara berurut-urut yaitu *input-ke-hidden, hidden-ke-output* dan *hidden-ke-hidden.* Contoh formula di atas merupakan RNN yang memiliki masukan *sequence*  dan keluaran *sequence* dengan Panjang yang sama (Goodfellow, Ian et al., 2017).

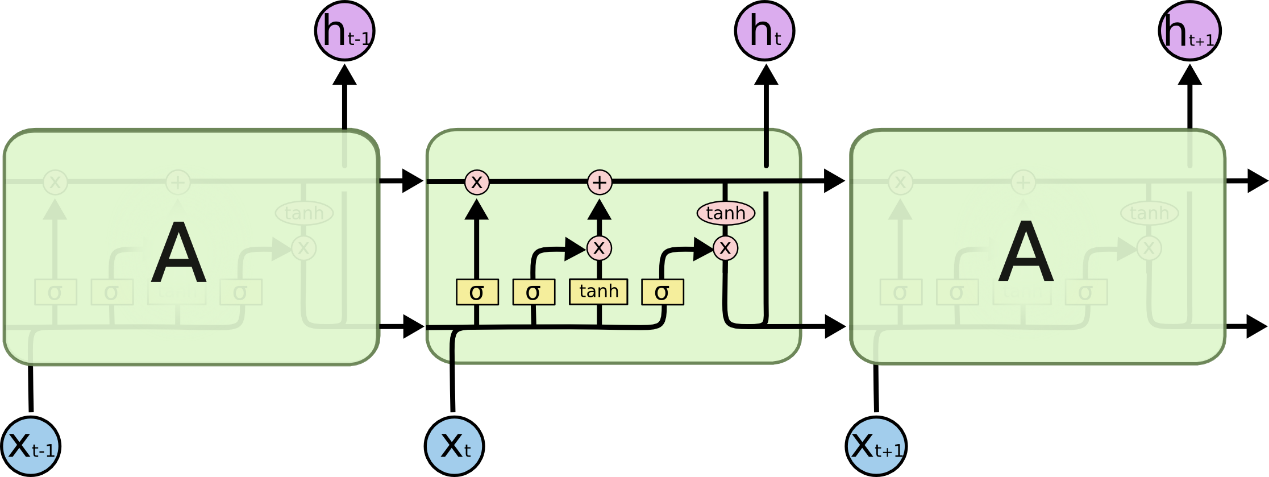
* + 1. **Backpropagation Trough Time**

*Backpropagation Through Time* merupakan *backpropagation* yang dilakukan pada RNN. *Backpropagation Through Time* atau disingkat dengan BPTT merupakan Teknik yang dilakukan untuk melatih RNN (Jan Wira Gotama Putra, et al., 2020). RNN tidak menggunakan *backprogation* untuk melakukan *training* dikarenakan metode tersebut kurang intuitif dalam menangani permasalahan data yang bersifat sekuensial (Jan Wira Gotama Putra, et al., 2020). Backpropagation melakukan pembaharuan parameter pada bobot dan bias sesuai error yang didapatkan pada saat foorward propagation dengan tujuan untuk meminimalkan error yang didapat.

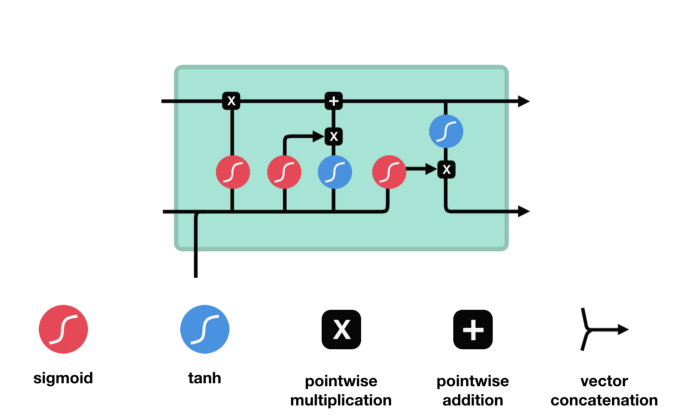
Pada backpropagation through time untuk memperbaharui parameter saat telah mencapai hidden state paling awal berbeda dengan *backpropagation* yang pada umumnya memperbaharui parameter bersamaan dengan mempropagasi error dari hidden state ke hidden state sebelumnya.. RNN merupakan salah satu neural network yang menggunakan teknik *stochastic gradient decent* yang merupakan variasi dari *gradient decent*. *Stohastic gradient decent* digunakan karena *gradient decent* tidak cocok dalam penggunaan dataset yang besar karena kemampuannya yang memakan waktu sangat lama. Dengan RNN yang menggunakan pencarian global mnimal error tersebut menghasilkan permasalahan yaitu *vanishing gradient* atau *exploiding gradient*. Kedua permasalahan tersebut terjadi karena pencarian parameter yang bisa mencapai nol (*vanisihing gradient)* atau melebihi batas seharusnya (*exploiding gradient)* sehingga dibutuhkan solusi lain dari hal tersebut salah satunya menggunakan metode baru yaitu LSTM.

### Long-Short Term Memory

Long Short Term Memory atau biasa disingkat dengan LSTM merupakan pengembangan dari metode RNN. LSTM didesain untuk menangani permasalahan pada saat pembelajaran *long-term dependencies.* Permasalahan tersebut bisa disebut juga sebagai *vanishing gradient* yaitu ketidakmampuan RNN dalam menangani menyimpan informasi pada jangka waktu yang panjang (long-term dependencies). Dengan LSTM hal tersebut bisa teratasi dengan *memory cell* dan *gate units.* Dengan *memory cell* menjadikan LSTM bisa memiliki kemampuan menyimpan dan menghapus informasi yang diregulasi oleh *gate units* (Sherstinsky, A., 2020).



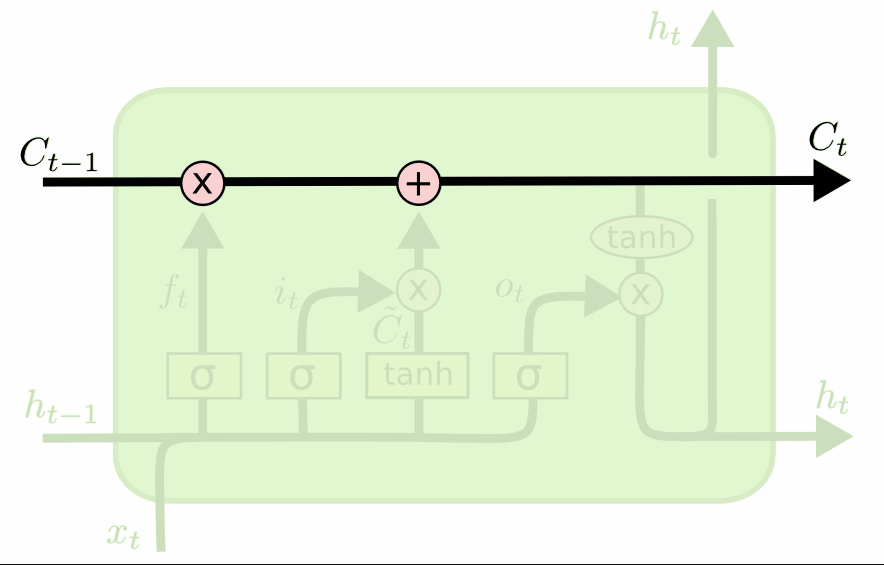
**Gambar 2.6 Diagram arsitektur LSTM**



**Gambar 2.7 Komponen pada LSTM**

Sesuai pada gambar 2.7 terdapat komponen – komponen pada LSTM yang membuat LSTM bekerja seperti fungsi *sigmoid* yang berfungsi sebagai memperbaharui atau mengubah nilai antara -1 dan 1 menjadi nilai yang diantara 0 dan 1, hal ini bertujuan untuk memperbaharui atau melupakan data karena angka yang dikalikan dengan angka 0 akan menghasilkan nilai 0 sehingga menyebabkan nilai menghilang dilain sisi setiap angka yang dikalikan dengan 1 akan tetap bernilai sama sehingga nilai tersebut akan tetap disimpan, sedangkan untuk fungsi *tanh* digunakan sebagai cara untuk mengontrol nilai yang melewati jaringan agar selalu berada diantara nilai -1 dan 1, operator pointwise serta *vector concatenation* sebagai jalannya proses LSTM. Pada gambar 2.6 terdapat simbol kotak yang berwarna kuning menunjukkan bahwa simbol tersebut adalah layer *neural network.*

Pada struktur jaringan LSTM tersusun oleh blok memori yang disebut dengan sel. Pada LSTM terdapat dua jalur yang menghubungkan informasi ke proses - proses selanjutnya yaitu *cell state* dan *hidden state*. Konsep inti pada LSTM adalah *cell state*. *Cell state* pada dasarnya merupakan "memory" pada jaringan LSTM. Informasi yang dibawa oleh cell state diatur oleh *gates* yang melakukan proses penambahan atau penghapusan informasi. Berikut pada gambar 2.8 merupakan struktur *cell state* pada LSTM.



**Gambar 2.8 Struktur cell state LSTM**

Pada gambar 2.8 diatas garis hitam lurus merupakan *cell state*.

### Attention based Mechanism

Attention adalah

* 1. **Transformer**

Transformer adalah

### 2.8 Bidirectional Encoder Representations from Transformers

BERT atau Bidirectional Encoder Representations from Transformers merupakan algoritma yang bertujuan untuk mempelajari representasi kata melalui dua tahapan tugas. Tugas pertama adalah memprediksi kata - kata yang hilang yang terdapat pada kalimat, berdasarkan sisa - sisa katanya yang terdapat pada kalimat tersebut. Tugas Kedua bertugas untuk menentukan apakah kalimat kedua dari dua kalimat adalah kalimat lanjutan atau apakah itu berasal dari teks yang sama sekali tidak berhubungan.

### 2.8 Pengujian

**2.8.1 Mean Average Precision**

**2.8.2 Root Mean Squared Error**

**2.9 Studi Pustaka (*State of the Art)***

Penelitian ini dilakukan karena...

|  |  |  |  |
| --- | --- | --- | --- |
| **No** | **Penulis** | **Judul** | **Rangkuman** |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |

# BAB III Metodologi Penelitian

## Metodologi Penelitian

## Teknik Pengumpulan Data

## Teknik Analisis Data

## Analisis Data

# BAB IV HASIL DAN PEMBAHASAN

# BAB V KESIMPULAN

# Daftar Pustaka

Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B,. (2011). Recommender Systems Handbook. In Recommender Systems Handbook. https://doi.org/10.1007/978-0-387-85820-3

Cheng, H. T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., Anil, R., Haque, Z., Hong, L., Jain, V., Liu, X., & Shah, H. (2016). Wide & deep learning for recommender systems. ACM International Conference Proceeding Series, 15-Septemb, 7–10. <https://doi.org/10.1145/2988450.2988454>

Burke, R., Felfernig, A., & Göker, M. H. (2011). Recommender Systems: An Overview. AI Magazine, 32(3), 13. doi:10.1609/aimag.v32i3.2361

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal, 16(3), 261–273. <https://doi.org/10.1016/j.eij.2015.06.005>

Modallal, S. (n.d.). Survey on Collaborative Filtering , Content-based Filtering and Hybrid Recommendation System Filtering and Hybrid Recommendation System (Modallal, S ., 2011).

Burke, R. (2007). Hybrid web recommender systems. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4321 LNCS, 377–408. <https://doi.org/10.1007/978-3-540-72079-9_12>

Suyanto., Nur Ramadhan, Kurniawan., Mandala, Satria., (2019) Penerbit Informatika

Alfarhood, M., & Cheng, J. (2021). Deep learning-based recommender systems. Advances in Intelligent Systems and Computing, 1232(1), 1–23. <https://doi.org/10.1007/978-981-15-6759-9_1>

Ng, Andrew. (n.d.). Structuring Machine Learning Projects [MOOC]. Coursera.

Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C. (2019). A review on deep learning for recommender systems: challenges and remedies. Artificial Intelligence Review, 52(1). https://doi.org/10.1007/s10462-018-9654-y

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., Gulcehre, C., Song, F., Ballard, A., Gilmer, J., Dahl, G., Vaswani, A., Allen, K., Nash, C., Langston, V., … Pascanu, R. (2018). Relational inductive biases, deep learning, and graph networks. 1–40. <http://arxiv.org/abs/1806.01261>

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D: Nonlinear Phenomena, 404(March), 1–43. <https://doi.org/10.1016/j.physd.2019.132306>

Goodfellow, Ian., Bengio, Yoshua., Courville, Aaron., (2016). Deep Learning. Nature, 26(7553), 436

Gao, Z., & Wang, X. (2019). Deep learning. EEG Signal Processing and Feature Extraction, 325–333. <https://doi.org/10.1007/978-981-13-9113-2_16>

Lops, P., Gemmis, M. De, & Semeraro, G. (2011). Recommender Systems Handbook. In Recommender Systems Handbook (Issue January). <https://doi.org/10.1007/978-0-387-85820-3>

Nuurshadieq, & Wibowo, A. T. (2020). Leveraging Side Information to Anime Recommender System using Deep learning. 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2020, 62–67. https://doi.org/10.1109/ISRITI51436.2020.9315363

Jindal, R., & Jain, K. (2019). A review on recommendation systems using deep learning. International Journal of Scientific and Technology Research, 8(10), 2978–2985.

Jan Wira Gotama Putra, et al. (2020). Pengenalan konsep pembelajaran mesin dan deep learning. Computational Linguistics and Natural Language Processing Laboratory, 4, 1–235. https://www.researchgate.net/publication/323700644

# Lampiran