# BAB II

## Tinjauan Pustaka

## Jenis Pembelajaran

Pembelajaran mesin adalah cabang dari kecerdasaran buatan yang memungkin sistem atau computer untuk belajar dari data yang tersedia yang kemudian mesin akan mempelajari untuk memprediksi sesuatu dari data tersebut. Dalam pembelajaran mesin, algoritma diajarkan untuk mengenali pola dalam data dan membuat keputusan berdasarkan pola tersebut. Dalam proses pembelajaran ada berbagai cara untuk melatih mesin tersebut, salah tiganya adalah:

## Supervised

Pembelajaran yang diawasi atau *supervised* adalah salah satu jenis pembelajaran mesin yang paling umum. Dalam pembelajaran terawasi, mesin mempelajari hubungan antara input dan output dari sampel data berlabel. Misalnya, dalam mengenali karakter tulisan tangan, mesin menerima gambar model karakter tulisan tangan dan label yang menunjukkan karakter mana yang ada di dalam gambar. Mesin kemudian menggunakan contoh-contoh ini untuk belajar mengenali pola yang terkait antara gambar dan karakter tulisan tangan. Mesin kemudian dapat mengenali karakter tulisan tangan dari gambar baru yang belum pernah terlihat sebelumnya. Pada dasarnya tujuan pembelajaran terbimbing adalah untuk menciptakan model atau fungsi matematika yang dapat menghubungkan input dengan output. Dalam hal ini, input adalah variabel atau atribut independen, sedangkan output adalah variabel dependen atau target. Tujuan dari pembelajaran terbimbing adalah untuk menemukan fungsi atau model yang dapat menghasilkan target yang akurat dan mendekati dunia nyata [1].

## Unsupervised

Pembelajaran tanpa pengawasan adalah jenis pembelajaran mesin yang berbeda dari pembelajaran yang diawasi. Dalam pembelajaran tanpa pengawasan, mesin tidak menerima sampel data berlabel. Sebagai gantinya, mesin harus belajar dari data yang tidak berlabel dan mencari kemungkinan pola dalam data tersebut. Pembelajaran tanpa pengawasan biasanya digunakan untuk mengelompokkan atau mengelompokkan data. Tujuannya adalah untuk mengelompokkan data yang serupa atau serupa ke dalam kelompok atau klaster tertentu. Contoh penggunaan pembelajaran tanpa pengawasan adalah dalam analisis data pasar, di mana kita dapat mengklasifikasikan konsumen dengan perilaku pembelian serupa ke dalam kelompok tertentu [1].

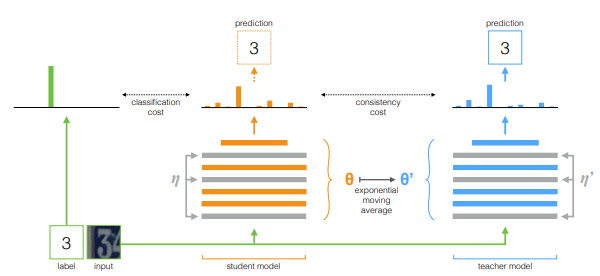
## Semi-Supervised

Pembelajaran semi-diawasi adalah teknik pembelajaran mesin yang mempelajari model dari data yang sebagian besar tidak berlabel dan hanya sebagian kecil dari data yang diberi label. Teknik ini berguna karena, dalam beberapa kasus, menemukan informasi yang diberi tag bisa jadi sulit dan mahal. Dalam pembelajaran semi-diawasi, model mencoba menemukan pola tersembunyi dalam data yang tidak berlabel sekaligus memanfaatkan informasi yang dikumpulkan dari data berlabel. Salah satu contoh penerapan semi-supervised learning adalah pengenalan wajah pada gambar, dimana hanya sebagian gambar saja yang diberi label sedangkan informasi lainnya tidak diberi label. Ada beberapa algoritma semi-supervised learning yang dapat digunakan, seperti *label propagation, joint training,* dan *self-training*. Algoritma ini bekerja dengan cara yang berbeda, tetapi semuanya bertujuan untuk menggunakan informasi dari data yang tidak berlabel untuk meningkatkan performa model [2].

## Jenis Algoritma

## Mean Teacher

*Mean Teacher* adalah salah satu algoritma dari jenis pembelajaran mesin *semi-supervised* dimana algoritma tersebut belajar dari data yang memiliki label dan tidak memiliki label. Algoritma ini berfungsi untuk melengkapi limitasi dari algoritma sebelumnya yaitu *Temporal Ensembling* [3], dimana algoritma *Temporal Ensembling* memperbarui targetnya satu kali per *epoch*, informasi yang dipelajari dimasukkan ke dalam proses pelatihan secara perlahan-lahan yang membuat semakin besar data yang dipelajari, semakin lama waktu yang dibutuhkan untuk pembelajaran. Algoritma *Mean Teacher* berusaha mengatasi limitasi tersebut dengan cara merata-ratakan bobot dari model dibandingkan dari hasil prediksi. Pada teknik *mean teachers*, model diajarkan pada data yang diberi label, dan juga pada data yang tidak diberi label. Model tersebut kemudian mengekstraksi fitur dari data yang tidak diberi label, dan memastikan konsistensi prediksi pada data yang sama oleh dua model yang berbeda, yang disebut *teacher* dan *student*. Teknik *weight-averaged consistency* target kemudian digunakan untuk memperbaiki konsistensi prediksi antara kedua model.

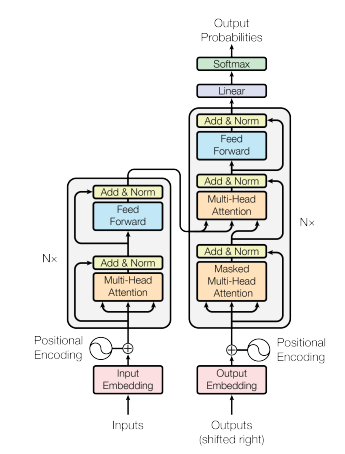
Algoritma *Mean Teacher* menunjukkan bahwa rata-rata *teachers* dan teknik target *weight-averaged consistency* yang konsisten dapat secara signifikan meningkatkan kinerja model dalam tugas pembelajaran *semi-supervised*, terutama untuk kumpulan data yang besar dan kompleks. Teknik ini juga dapat digunakan dengan augmentasi data input dan teknik denoising untuk meningkatkan kinerja model. Hasil percobaan pada beberapa dataset menunjukkan bahwa penggunaan teknik ini dapat meningkatkan akurasi model secara signifikan dibandingkan dengan teknik pembelajaran *semi-supervised* yang hanya menggunakan teknik augmentasi atau *denoising* data hanya pada input data [4].

**Figure 1 Alur dari Algoritma Mean Teacher**

## Transformer

Algoritma transformer adalah sebuah arsitektur *deep learning* yang hanya menggunakan *Self-Attention* didalamnya tanpa menggunakan algoritma konvolusi maupun *reccurent*. Transformer menggunakan blok transformer yang terdiri dari *encoder* dan *decoder*. *Encoder* bertanggung jawab untuk mempelajari representasi kalimat atau data dalam bentuk *embedding*, sedangkan *decoder* bertanggung jawab untuk memprediksi keluaran terjemahan dalam proses pembelajaran translasi bahasa. *Attention* digunakan di setiap blok *encoder* dan *decoder* untuk menghubungkan setiap bagian dalam blok tersebut dari kalimat *input* dengan setiap bagian kalimat *output*. Hal tersebut yang memungkinkan Transformer untuk memperhatikan bagian mana dari kalimat tersebut yang relevan ketika proses pembelajaran [5].

Berbeda dengan algoritma pada umumnya yang menggunakan layar konvolusi ataupun *reccurent*, Transformer hanya menggunakan algoritma *Self-Attention* didalamnya. *Self-Attention* berfungsi untuk memberikan bobot pada seluruh *input* data untuk membangun representasi yang lebih baik, cara kerja dari *Self-Attention* sebagai berikut:

* + - 1. Pembuatan *Query, Key,* dan *Value*, pada tahapain ini, setiap kata pada *input sequences* akan diproyeksikan kedalam tiga vektor Q, K, dan V dengan dimensi yang sama
      2. Perhitungan bobot, selanjutnya, bobot akan dihitung dengan melakukan *dot-product* antara Q dengan K, kemudian hasil tersebut akan dibagi dengan akar kuadrat dari dimensi K. Hasil dari perhitungan tersebut kemudian akan dimasukkan kedalam fungsi *Softmax* untuk menghasilkan bobot yang menentuikan seberapa pentingnya suatu *Key* dalam membangun representasi *Query.*
      3. Perkalian bobot representasi dengan *Value*, setelah bobot representasi dihitung, bobot tersebut akan dikalikan secara *dot-product* dengan V untuk menghasilkan representasi yang lebih baik. Hasil dari perhitungan tersebut adalah *output* dari *Self-Attention* yang kemudian dilanjutkan ke layer berikutnya.

**Figure 2 Alur dari Algoritma Transformer**

## 

## Loss Function

*Loss Function* adalah suatu metrik yang digunakan untuk mengevaluasi seberapa baik suatu model pembelajaran mesin dapat memepelajari suatu pola dari data yang diberikan. *loss function* mengukur perbedaan antara nilai prediksi model dengan nilai target yang seharusnya, dan mencoba untuk meminimalkan perbedaan ini selama proses pembelajaran. *loss function* dapat digunakan dalam berbagai tugas pembelajaran mesin, seperti klasifikasi, regresi, segmentasi, dan rekonstruksi. Masing-masing tugas membutuhkan jenis *loss function* yang berbeda-beda tergantung pada sifat data dan target yang ingin dicapai [6].

## Classification Loss

*Classification Loss* adalah suatu metrik yang digunakan dalam proses pembelajaran mesin *supervised* untuk mengevaluasi seberapa baik model dapat mengklasifikasikan input ke dalam beberapa kategori yang tepat. Metrik ini mengukur seberapa jauh antara prediksi model dan label sebenernya dari input dan berusaha untuk meminimalkan perbedaan ini selama pembelajaran [6]. Terdapat beberapa jenis *classification loss* yang dapat digunakan seperti *Cross Entropy Loss, Hinge Loss, Squared Hinge Loss*, dan lain lain. Masing-masing jenis *loss* ini memiliki kegunaan dan karakteristik yang berbeda tergantung pada tugas yang sedang dijalankan dan jenis model yang digunakan.

## Consistency Loss

*Consistency loss* adalah suatu metrik yang digunakan dalam proses pembelajaran mesin *semi-supervised* untuk meningkatkan konsistensi prediksi model terhadap data yang sama Dilihat dari sudut pandang yang berbeda-beda. *Consistency loss* digunakan untuk memastikan bahwa model memberikan prediksi yang konsisten ketika data yang sama diberikan dalam bentuk yang berbeda-beda seperti contoh, jika suatu model memberikan prediksi yang berbeda-beda ketika data diberikan dalam bentuk asli dan bentuk yang sudah dimodifikasi, maka *consistency loss* akan digunakan untuk mengatasi permasalahan tersebut. *Consistency loss* sangat penting dalam pembelajaran mesin *semi-supervised* dimana *loss* tersebut akan berfungsi untuk meningkatkan konsistensi prediksi model [6].

## Ringkasan Literatur

*Topic* *Modeling* adalah suatu metode dalam analisis teks yang digunakan untuk mengidentifikasi dan mengelompokkan topik-topik yang terdapat dalam suatu kumpulan teks tanpa adanya pengawasan manusia (*Unsupervised Learning*) metode ini menggunakan teknik statistik untuk mengidentifikasi pola-pola tematik yang muncul dalam *teks*, seperti kata-kata yang sering muncul bersama-sama atau pola sintaksis yang sering digunakan. Salah satu algoritma yang umum digunakan dalam *topic modelling* adalah *Latent Dirichlet Allocation* (LDA), yang menganggap bahwa setiap dokumen dalam kumpulan teks mengandung beberapa topik tersembunyi yang sama dengan dokumen lainnya. Dengan menggunakan LDA, *topic modelling* dapat menghasilkan kumpulan topik topik yang saling berkaitan berdasarkan distribusi probabilitas kata-kata dalam dokumen-dokumen yang dianalis.

Kajian literatur merupakan salah satu langkah penting dalam proses penelitian untuk mengidentifikasi, mengkaji, dan menyintesis pengetahuan yang telah ada dalam suatu bidang penelitian. Dalam era informasi digital saat ini, jumlah publikasi ilmiah yang tersedia terus meningkat dengan cepat, sehingga menyulitkan peneliti untuk menghadapinya. Oleh karena itu, metode analisis menggunakan LDA telah menjadi alat yang sangat berguna dalam membantu melakukan kajian ilmiah dengan cara yang lebih efisien dan efektif.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Image Classification | | Algorithms for Problem Solving | | Social Media Analysis | | Natural Language Processing | | Machine Learning | |
| Tema | Distribusi Marginal | Tema | Distribusi Marginal | Tema | Distribusi Marginal | Tema | Distribusi Marginal | Tema | Distribusi Marginal |
| Image | 0.019 | Problem | 0.012 | Social | 0.017 | Language | 0.011 | Feature | 0.013 |
| Task | 0.014 | Algorithms | 0.010 | Event | 0.014 | System | 0.011 | Prediction | 0.009 |
| Classification | 0.008 | Task | 0.008 | Medium | 0.014 | Application | 0.009 | Neural | 0.008 |
| Dataset | 0.008 | Approach | 0.007 | Information | 0.013 | Processing | 0.009 | Task | 0.008 |
| Approach | 0.007 | Agent | 0.007 | Twitter | 0.010 | Research | 0.008 | Approach | 0.007 |
| Language | 0.007 | System | 0.006 | Tweet | 0.009 | Intelligence | 0.008 | Layer | 0.006 |
| Label | 0.007 | Training | 0.006 | Time | 0.007 | Neural | 0.007 | Algorithms | 0.006 |
| Training | 0.007 | Performance | 0.006 | Study | 0.007 | Artificial | 0.006 | High | 0.005 |
| Datasets | 0.006 | Reinforcement | 0.005 | Flood | 0.006 | Challenge | 0.006 | Performance | 0.005 |
| Performance | 0.006 | Framework | 0.005 | People | 0.006 | Provide | 0.006 | Cluster | 0.005 |

## Penelitian Terkait

Beberapa penelitian terkait hasil dari LDA yang sudah pernah dilakukan dan telah diterbitkan antara lain:

1. Gonzalo Mateo-Garcia et al. [7] dalam penelitian mereka yang berjudul “*Flood Detection On Low Cost Orbital Hardware*” mengumpulkan data gambar banjir dari satelit yang tersedia secara publik kemudian menggunakan *Convolutional Neural Network* untuk melakukan prediksi dari banjir dan melakukan segmentasi pada gambar tersebut untuk mendeteksi banjir. Hasil dari penelitian tersebut menggunakan arsitektur U-Net mendapatkan hasil dari metrik IOU sebesar 69.25% pada jarak 10 meter, dan 67.51% pada arsitektur U-Net untuk metrik IOU pada jarak 80 meter. Dan untuk metrik Recall pada jarak 10 meter arsitektur SCNN mendapatkan skor tertinggi dengan 97.43% dan pada jarak 80 meter arsitektur SCNN juga mendapatkan skor tertinggi dengan 97.03%. mereka menunjukkan bahwa melakukan segmentasi dan prediksi banjir mampu dilakukan menggunakan arsitektur *deep learning* meskipun dilakukan pada gambar yang beresolusi rendah yang didapat dari satelit.
2. Maryam Rahnemoonfar et al. [8] dalam penelitian mereka yang berjudul “*Comprehensive Sementic Segmentation on High Resolution UAV Imagery for Natural Disaster Damage Assessmentí*” melakukan segmentasi dari data pemandangan yang diambil oleh satelit untuk mengklasifikasi apakah daerah tersebut tidak rusak, rusak ringan, rusak berat, maupun hancur total. Hasil penelitian tersebut mendapatkan bahwa karna gambar yang didapat berdasarkan dari satelit, hanya kerusakan pada atap yang sangat mempengaruhi model untuk memprediksi pada kelas apa gambar tersebut dikategorikan, **meskipun begitu** PSPNet dapat mengklasifikasi gambar tersebut dengan baik mengalahkan arsitektur yang lain dan rata-rata IOU sebesar 99.79%
3. Hitoshi Shimizu et al. [9] dalam penelitian mereka yang berjudul “*Evacuation Shelter Scheduling Problem*” berusaha mengatasi permasalahan utama yang terjadi ketika terjadi bencana alam yakni ketika terjadi bencana alam yang besar maka dibutuhkan juga dana yang besar untuk mengevakuasi korban bencana. Dan ketika memindahkan lokasi tempat evakuasi juga tidak bisa dilakukan **sembarangan**, karna akan berdampak pada biaya yang terbuang dan juga kondisi penghuni lokasi tempat evakuasi. Hasil penelitian tersebut mendapatkan dengan menggunakan algoritma BinPack dapat menghemat biaya evakuasi sebesar 25.4 juta USD.
4. Guotai Wang et al. [10] dalam penelitian mereka yang berjudul “A Noise-Robust Framework for Automatic Segmentation of COVID-19 Pneumonia Lesions From CT Images” mencoba mengatasi permasalahan dimana susahnya mendapatkan data gambar CT yang berkualitas tinggi. Dengan hanya tersedia data gambar CT yang memiliki kualitas rendah, **Guotai Wang et al.** menggunakan pendekatan algoritma *mean teacher* untuk membuat model *noise-robust* yang dapat mempelajari dari gambar CT berkualitas rendah. Hasil penelitian tersebut berhasil meningkatkan akurasi dengan memanfaatkan *mean teacher* dengan perbandingan Attention U-Net tanpa menggunakan *mean teacher* menghasilkan *dice* metrik sebesar 77.23% sedangkan COPLE-Net yang dilatih menggunakan *mean teacher* memiliki *dice* metrik 80.29%. dari hasil tersebut **Guotai Wang et al.**  menyimpulkan bahwa dengan menggunakan algoritma *mean teacher* dapat membantu model dimana data yang memiliki label sedikit dan susah didapat selain itu data gambar memiliki kualitas yang rendah.
5. Arnab Roy, Muneendra Ojha [11] dalam penelitian mereka yang berjudul “Twitter Sentiment Analysis Using Deep Learning” membandingkan arsitektur *convolutional neural network*, *pretrained* glove, *Bidirectional-Attention* LSTM, dan BERT untuk **mengklasfikasi** sentiment dari data Twitter. Hasil penelitian mereka menemukan bahwa BERT berhasil mengalahkan semua arsitektur lainnya dan mendapatkan metrik F1 Score sebesar 63% sedangkan arsitektur *Bidirectional-Attention* LSTM mendapatkan F1 Score sebesar 61%.
6. Artaches Ambartsoumian, Fred Popowich [12] dalam penelitian mereka yang berjudul “Self-Attention: A Better Building Block for Sentiment Analysis Neural Network Classifier” **ingin** membandingkan bagaimana Transformer dengan tradisional *neural network* seperti convolutional, LSTM, dan *Bidirectional*-LSTM dibeberapa dataset untuk pengujian. Hasil penelitian mereka menemukan bahwa transformer mampu menghasilkan akurasi yang lebih tinggi dibandingkan dengan tradisional *neural network* disemua dataset.

[1] Buku “An Introduction to Machine Learning”, Alpaydin E. (2010)

[2] Buku “Machine Learning: A Probabilistic Approach”, Murphy K. P. (2012)

[3] Paper “Temporal Ensembling for Semi-Supervised Learning” (2016)

[4] Paper “Mean Teacher are better role models” (2018)

[5] Paper “Attention is All You Need” (2017)

[6] Buku “Deep Learning” Goodfellow, Bengio (2016)

[7] [1910.03019v3.pdf (arxiv.org)](https://arxiv.org/pdf/1910.03019v3.pdf)

[8] [2009.01193v2.pdf (arxiv.org)](https://arxiv.org/pdf/2009.01193v2.pdf)

[9] [2111.13326v1.pdf (arxiv.org)](https://arxiv.org/pdf/2111.13326v1.pdf)

[10] <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9109297>

[11] <https://sci-hub.se/https://ieeexplore.ieee.org/abstract/document/9342279>

[12] https://arxiv.org/pdf/1812.07860.pdf