

Autoencoders

Ali Akbar Septiandri

Universitas Al Azhar Indonesia

December 9, 2018

Daftar isi

1. Autoencoders

2. Embeddings

Representasi laten

Word2vec

Bahan Bacaan

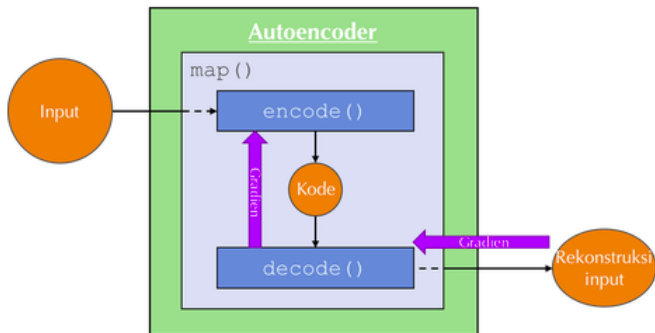
1. Goodfellow, I., et al. 2016. Deep Learning. MIT Press. **Chapter 14: Autoencoders.**
2. Chollet, F. 2016. Building Autoencoders in Keras. URL: <https://blog.keras.io/building-autoencoders-in-keras.html>
3. Wibisono, O. 2017. Autoencoder: Alternatif PCA yang Lebih Mumpuni. URL: <https://tentangdata.wordpress.com/2017/10/14/autoencoder-alternatif-pca-yang-lebih-mumpuni/>

Autoencoders

Autoencoders

- Arsitektur *neural networks* yang memetakan inputnya ke dalam suatu **representasi** yang lebih **sederhana** yang digunakan untuk **merekonstruksi inputnya kembali**
- Diperkenalkan dalam [Hinton and Salakhutdinov, 2006]
- Rekonstruksi yang dihasilkan bersifat **lossy**
- Digunakan untuk **mereduksi dimensi**

Ilustrasi



Gambar: Arsitektur umum autoencoder [Wibisono, 2017]

Contoh Kode

```
input_img = Input(shape=(784,))
encoded = Dense(128, activation='relu')(input_img)
encoded = Dense(64, activation='relu')(encoded)
encoded = Dense(32, activation='relu')(encoded)

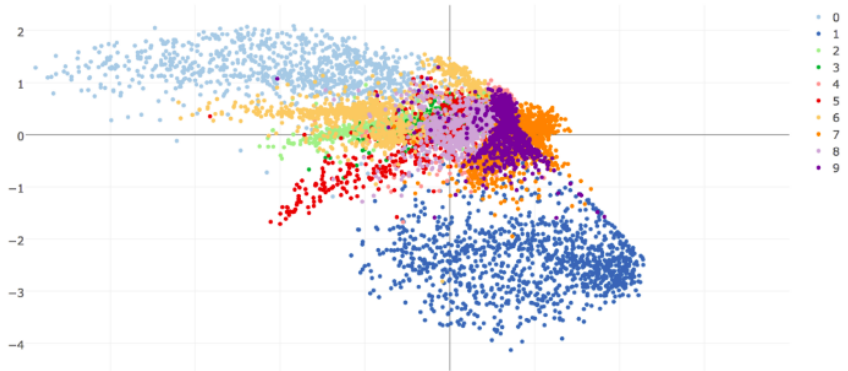
decoded = Dense(64, activation='relu')(encoded)
decoded = Dense(128, activation='relu')(decoded)
decoded = Dense(784, activation='sigmoid')(decoded)

autoencoder = Model(input_img, decoded)
autoencoder.compile(optimizer='adadelta', loss='binary_crossentropy')

autoencoder.fit(x_train, x_train,
                epochs=100,
                batch_size=256,
                shuffle=True,
                validation_data=(x_test, x_test))
```

Gambar: Diadaptasi dari [Chollet, 2016]

Pemanfaatan



Gambar: Pengelompokan digit dengan autoencoder [Wibisono, 2017]

Denoising Autoencoders

- Ide: Memetakan input yang rusak menjadi yang bersih

Denoising Autoencoders

- Ide: Memetakan input yang rusak menjadi yang bersih
- Sifat autoencoder yang *lossy* memungkinkan ini, i.e. membaca sinyal penting saja

Denoising Autoencoders

- Ide: Memetakan input yang rusak menjadi yang bersih
- Sifat autoencoder yang *lossy* memungkinkan ini, i.e. membaca sinyal penting saja
- Alih-alih menggunakan *loss function*

$$L(\mathbf{x}, g(f(\mathbf{x}))),$$

fungsi yang digunakan adalah

$$L(\mathbf{x}, g(f(\tilde{\mathbf{x}})))$$

Contoh Kasus

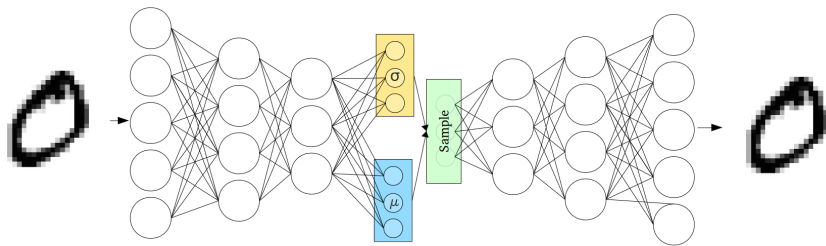


Gambar: Membuang noise dari gambar dengan autoencoder [Chollet, 2016]

Variational Autoencoders

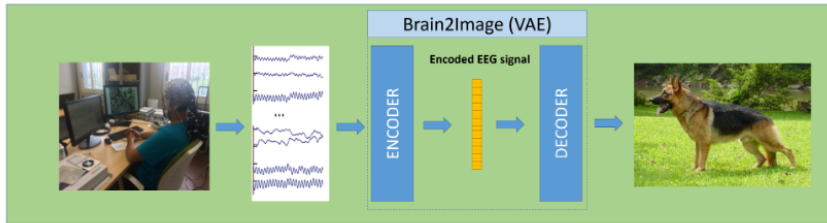
- *Latent spaces*-nya berupa suatu *distribusi* sehingga memungkinkan *sampling*
- VAE bersifat *generatif*, i.e. dapat menghasilkan data baru
- Dikenalkan pertama kali dalam [Kingma and Welling, 2013]

Arsitektur



Gambar: Contoh Variational Autoencoder (VAE) [Shafkat, 2018]

Pemanfaatan



Gambar: Mencoba memetakan sinyal EEG untuk merekonstruksi gambar [Kavasidis et al., 2017]

Embeddings

Representasi Laten

- Ide merepresentasikan data dalam dimensi yang lebih rendah bisa diaplikasikan ke jenis data lain

Representasi Laten

- Ide merepresentasikan **data dalam dimensi yang lebih rendah** bisa diaplikasikan ke jenis data lain
- Digunakan juga untuk:
 - sistem rekomendasi dengan *collaborative filtering* [Salakhutdinov et al., 2007]
 - representasi kata [Mikolov et al., 2013]
 - *image captioning* [Vinyals et al., 2015]

Representasi Kata

- Ide merepresentasikan data dalam dimensi yang lebih rendah bisa diaplikasikan ke jenis data lain

Representasi Kata

- Ide merepresentasikan **data dalam dimensi yang lebih rendah** bisa diaplikasikan ke jenis data lain
- **Dokumen** biasanya direpresentasikan hanya dalam *bag-of-words*

Representasi Kata

- Ide merepresentasikan data dalam dimensi yang lebih rendah bisa diaplikasikan ke jenis data lain
- Dokumen biasanya direpresentasikan hanya dalam *bag-of-words*
- Apakah bisa dibuat hubungan antarkata?

Contoh

D1: “send us your password” (s)

D2: “send us your review” (h)

D3: “review your password” (h)

D4: “review us” (s)

D5: “send your password” (s)

D6: “send us your account” (s)

Contoh (lanjutan)

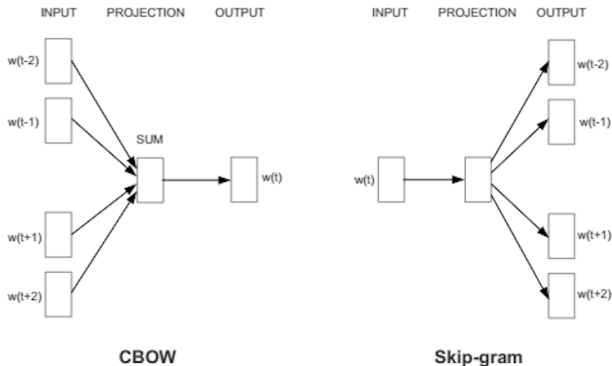
Dengan menggunakan one-hot-encoder

account	password	review	send	us	your	y
0	1	0	1	1	1	1
0	0	1	1	1	1	0
0	1	1	0	0	1	0
0	0	1	0	1	0	1
0	1	0	1	0	1	1
1	0	0	1	1	1	1

word2vec

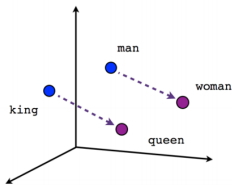
- Representasi laten dapat dihasilkan dengan **word2vec** [Mikolov et al., 2013]
- Menggunakan *neural networks* yang terdiri dari **dua layers**
- Dapat menggunakan dua jenis arsitektur: **CBOW** dan **skip-gram**

Arsitektur

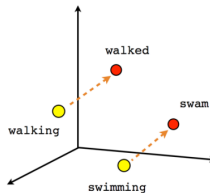


Gambar: Arsitektur CBOW dan skip-gram untuk word2vec
[Mikolov et al., 2013]

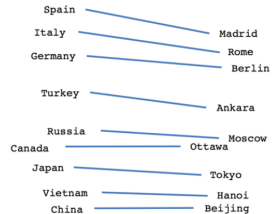
Contoh Hasil



Male-Female



Verb tense



Country-Capital

Gambar: Ilustrasi dari representasi yang dihasilkan [TensorFlow, 2018]

Penggunaan

- Representasi tersebut memungkinkan persamaan

$$\textit{king} - \textit{queen} = \textit{man} - \textit{woman}$$

- Dapat menggunakan pustaka gensim
- Contoh aplikasi di Rare Technologies

Perkembangan

- Sudah banyak perkembangan algoritma, e.g. doc2vec, FastText

Perkembangan

- Sudah banyak perkembangan algoritma, e.g. doc2vec, FastText
- *Language modelling* → *transfer learning*

Perkembangan

- Sudah banyak perkembangan algoritma, e.g. doc2vec, FastText
- *Language modelling* → *transfer learning*
- Melatih model semudah penggunaan ResNet untuk gambar, dimulai dari ELMo dan ULMFit

Perkembangan

- Sudah banyak perkembangan algoritma, e.g. doc2vec, FastText
- *Language modelling* → *transfer learning*
- Melatih model semudah penggunaan ResNet untuk gambar, dimulai dari ELMo dan ULMFit
- Tahun 2018, Google mengeluarkan BERT

Ikhtisar

1. *Neural networks* tidak hanya memprediksi, tapi juga **mengekstraksi fitur**
2. *Autoencoders* adalah metode *unsupervised learning* untuk **reduksi dimensi** → representasi data
3. VAE dapat digunakan untuk menghasilkan data baru dengan **sampling**
4. Untuk representasi kata, dapat menggunakan **word2vec**
5. Bidang yang sangat terbuka untuk penelitian karena sifatnya yang **spesifik bahasa**

Referensi



Geoffrey Hinton and Ruslan Salakhutdinov (2006)

Reducing the dimensionality of data with neural networks
Science, 313(5786), pp.504-507



Okiriza Wibisono (October 2017)

Autoencoder: Alternatif PCA yang Lebih Mumpuni

[https://tentangdata.wordpress.com/2017/10/14/
autoencoder-alternatif-pca-yang-lebih-mumpuni/](https://tentangdata.wordpress.com/2017/10/14/autoencoder-alternatif-pca-yang-lebih-mumpuni/)



Francois Chollet (2016)

Building Autoencoders in Keras

<https://blog.keras.io/building-autoencoders-in-keras.html>

Referensi



Diederik Kingma and Max Welling (2013)

Auto-encoding variational bayes

arXiv preprint arXiv:1312.6114



Irhum Shafkat (2018)

Intuitively Understanding Variational Autoencoders

[https://towardsdatascience.com/](https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf)

[intuitively-understanding-variational-autoencoders-1bfe67eb5daf](https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf)







Kavasidis, I., Palazzo, S., Spampinato, C., Giordano, D. and Shah, M. (2017)

Brain2Image: Converting Brain Signals into Images

In Proceedings of the 2017 ACM on Multimedia Conference (pp. 1809-1817).
ACM.

Referensi

-  Salakhutdinov, R., Mnih, A. and Hinton, G. (2007)
Restricted Boltzmann machines for collaborative filtering
In Proceedings of the 24th ICML (pp. 791-798). ACM.
-  Vinyals, O., Toshev, A., Bengio, S. and Erhan, D. (2015)
Show and tell: A neural image caption generator
In CVPR (pp. 3156-3164)
-  Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013)
Distributed representations of words and phrases and their compositionality
In NIPS (pp. 3111-3119)
-  TensorFlow (Last updated November 20, 2018)
Vector Representations of Words
<https://www.tensorflow.org/tutorials/representation/word2vec>

Terima kasih