

EI7007 Pebelejaran Mesin Lanjut

Tugas Eksplorasi Hyper Parameter

Nama : Ahmad Luky Ramdani

NIM: 33221020



**Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung
Bandung**

Pada bagian ini saya akan mencoba melakukan eksplorasi terkait dengan proses optimasi hyperparameter yang terdapat dalam arsitektur CNN dan MLP. Pada eksplorasi ini saya akan mencoba menggunakan 2 data dengan 2 jenis persoalan yaitu:

- Persoalan klasifikasi dengan menggunakan data Fashion MNIST
- Persoalan regresi dengan menggunakan data Boston Housing Price

A. Persoalan Klasifikasi untuk arsitektur CNN (Dataset Fashion MNIST)

Sehingga setelah proses eksplorasi didapatkan informasi terkait dengan hyperparameter pada model CNN yang didapatkan. Hyperparameter tersebut seperti:

1. Jumlah convolution layer yang optimal
2. Ukuran filter yang optimal untuk setiap convolution layer
3. Banyaknya filter yang optimal untuk setiap convolution layer
4. Banyaknya hidden unit yang optimal pada bagian fully connected network

Selain hyperparameter di atas, dilakukan eksplorasi terkait dengan **Optimizer**, **learning rate schedule** dan **Losses** pada nilai parameter default untuk mendapatkan kinerja model paling baik

Proses eksplorasi dilakukan tahapan

1. Menggunakan 1 layer ConV
2. Menggunakan 2 layer ConV

Pada eksplorasi menggunakan **1 layer ConV** dengan melihat perubahan jumlah hidden unit, didapatkan informasi nilai akurasi sbb

Conv Layer Ke-	Jumlah Hidden Unit	Jumlah Filter	Ukuran Filter	Data Training		Data Testing	
				Akurasi	Loss	Akurasi	Loss
1	32	128	3x3	0.9325	0.1862	0.9077	0.2622
1	32	128	5x5	0.9263	0.2022	0.9061	0.2627
1	64	128	3x3	0.9348	0.1774	0.9051	0.2695
1	64	128	5x5	0.9333	0.1803	0.9118	0.2507
1	128	128	3x3	0.9428	0.1552	0.9124	0.2624
1	128	128	5x5	0.9394	0.1645	0.9117	0.2679

Dari tabel di atas nilai maksimal akurasi yang didapatkan dari kombinasi

1. Jumlah filter = 128
2. Ukuran filter = 3 x3
3. Jumlah Hidden Unit = 128

Selanjutnya dilakukan proses eksplorasi dengan menambahkan 2 layer ConV

No.	Conv Layer Ke-	Jumlah Hidden Unit	Jumlah Filter	Ukuran Filter	Data Training		Data Testing	
					Akurasi	Loss	Akurasi	Loss
1	2	32	128	3x3	0.9235	0.2061	0.9087	0.2509
2	2	32	128	5x5	0.926	0.2029	0.9063	0.2588
3	2	32	128	7x7	0.9255	0.2001	0.9046	0.2683
4	2	32	128	9x9	0.9298	0.1922	0.8926	0.297
5	2	64	128	3x3	0.9256	0.2025	0.9037	0.2558
6	2	64	128	5x5	0.9306	0.1872	0.9036	0.2633
7	2	64	128	7x7	0.9251	0.203	0.906	0.2647
8	2	64	128	9x9	0.9323	0.1833	0.9	0.2765
9	2	128	128	3x3	0.9295	0.1846	0.901	0.2728
10	2	128	128	5x5	0.9328	0.1835	0.9073	0.2685
11	2	128	128	7x7	0.932	0.1851	0.9058	0.275
12	2	128	128	9x9	0.9374	0.1699	0.9034	0.2787

Dari tabel di atas nilai maksimal akurasi yang didapatkan dari kombinasi

1. Jumlah filter = 128
2. Ukuran filter = 9 x 9
3. Jumlah Hidden Unit = 128

B. Persoalan Regresi untuk arsitektur MLP (Data Boston Housing Price)

Kasus ini menggunakan persoalan pada regresi dengan menggunakan arsitektur Neural Network. Sehingga setelah proses eksplorasi didapatkan informasi terkait dengan hyperparameter yaitu:

1. Jumlah hidden layer yang optimal
2. Jumlah hidden unit yang optimal di setiap hidden layer
3. Activation function dengan hasil optimal
4. Optimizer yang hasil optimal
5. Loss function dengan hasil optimal

Proses eksplorasi dilakukan dengan melakukan percobaan pada setiap hyperparamter. Berikut hasil eksplorasi terkait dengan hyperparameter

No.	Jumlah Hidden Layer	Jumlah Hidden Unit	Activation Function	Optimizer	Rata-rata MSE
1	1	13	relu	Adam	33.93
2	1	13	relu	SGD	6.94
3	1	13	relu	RMSprop	33.64

4	1	13	relu	SGD	6.74
5	1	13	softmax	SGD	12.84
6	1	32	relu	SGD	5.17
7	1	64	relu	SGD	4.85
8	1	128	relu	SGD	3.93
9	2	64 (Layer 1), 32 (Layer 2)	relu	SGD	3.41
10	2	64 (Layer 1), 64 (Layer 2)	relu	SGD	2.58
11	2	64 (Layer 1), 128 (Layer 2)	relu	SGD	2.48
12	2	64 (Layer 1), 128 (Layer 2)	softmax	SGD	8.88
13	2	128 (Layer 1), 32 (Layer 2)	relu	SGD	3.46
14	2	128 (Layer 1), 64 (Layer 2)	relu	SGD	2.15
15	2	128 (Layer 1), 128 (Layer 2)	relu	SGD	2.24
16	2	128 (Layer 1), 128 (Layer 2)	softmax	SGD	10.47

Dari hasil eksplorasi hyperparameter di atas didapatkan model dengan hyperparameter yang menghasilkan nilai MSE terkecil adalah

1. Jumlah Layer = 2
2. Jumlah Hidden Unit Layer 1 = 128
3. Jumlah Hidden Unit Layer 2 = 64
4. Activation Function Layer 1 = ReLu
5. Activation Function Layer 2 = ReLu
6. Optimizer = SGD
7. Loss Function = MSE