

1. Hidden Size

Output menunjukkan bahwa model dilatih dengan tiga ukuran hidden size: **16**, **32**, dan **64**. Berikut adalah hasil analisis:

Hidden Size	Final Train Loss	Final Validation Loss
16	0.3927	0.3924
32	0.3930	0.3922
64	0.3934	0.3917

- **Hidden Size 16:**
 - Train Loss mencapai nilai terendah dibanding hidden size lainnya, yaitu **0.3927**.
 - Namun, Validation Loss sedikit lebih tinggi dibanding hidden size 64 (**0.3924** vs. **0.3917**).
 - Hidden size ini cenderung overfitting pada data training, terlihat dari perbedaan yang lebih kecil antara Train Loss dan Validation Loss.
- **Hidden Size 32:**
 - Train Loss dan Validation Loss cukup konsisten (**0.3930** dan **0.3922**).
 - Hidden size ini menunjukkan performa yang stabil, cocok untuk model yang tidak terlalu kompleks dan mudah diatur.
- **Hidden Size 64:**
 - Meskipun Train Loss tertinggi (**0.3934**), Validation Loss terendah (**0.3917**).
 - Hidden size ini memberikan generalisasi yang lebih baik pada data validasi.
 - Ukuran hidden size yang lebih besar membantu menangkap lebih banyak fitur tetapi meningkatkan risiko overfitting jika epoch diperpanjang.

Kesimpulan untuk Hidden Size: Hidden size **64** menghasilkan performa terbaik pada Validation Loss. Ini menunjukkan bahwa kapasitas model yang lebih besar membantu menangkap pola yang lebih kompleks pada data validasi.

2. Analisis Berdasarkan Optimizer

Model dilatih menggunakan tiga optimizer: **SGD**, **RMSProp**, dan **Adam**. Berikut adalah hasil analisis:

Optimizer	Final Train Loss	Final Validation Loss
SGD	0.3998	0.3970
RMSProp	0.3934	0.3913

Optimizer	Final Train Loss	Final Validation Loss
Adam	0.3930	0.3910

- **SGD (Stochastic Gradient Descent):**

- Train Loss tertinggi (**0.3998**) dan Validation Loss tertinggi (**0.3970**) dibandingkan optimizer lainnya.
- SGD lebih lambat untuk konvergen karena gradien di-update secara langsung tanpa adaptasi.

- **RMSProp:**

- Train Loss dan Validation Loss cukup rendah (**0.3934** dan **0.3913**).
- RMSProp memberikan stabilitas lebih baik dibanding SGD dengan menangani learning rate yang adaptif.

- **Adam:**

- Train Loss dan Validation Loss terendah (**0.3930** dan **0.3910**).
- Adam menunjukkan performa terbaik untuk pelatihan ini karena mampu mengadaptasi learning rate dengan baik pada setiap parameter.

Kesimpulan untuk Optimizer: Optimizer **Adam** memberikan performa terbaik, menghasilkan Train Loss dan Validation Loss yang paling rendah. Ini menunjukkan bahwa Adam lebih cocok untuk menangani data dengan pola yang kompleks dibandingkan SGD dan RMSProp.

3. Analisis Berdasarkan Jumlah Epoch

Model dilatih dengan dua jumlah epoch: **5** dan **50**. Berikut adalah hasil analisis:

Epoch	Final Train Loss	Final Validation Loss
5	0.3944	0.3916
50	0.3921	0.3914

- **Epoch 5:**

- Train Loss lebih tinggi (**0.3944**) karena model belum cukup lama dilatih untuk mempelajari pola data sepenuhnya.
- Validation Loss sedikit lebih tinggi (**0.3916**) dibandingkan epoch 50.

- **Epoch 50:**

- Train Loss menurun menjadi **0.3921**, menunjukkan model mempelajari data lebih baik.
- Validation Loss mencapai **0.3914** tetapi tidak banyak berubah dibandingkan epoch 5.

- Early stopping diaktifkan setelah 17 epoch, menunjukkan bahwa pelatihan tambahan setelah titik ini tidak memberikan keuntungan signifikan.

Kesimpulan untuk Epoch: Jumlah epoch **50** memberikan performa terbaik, tetapi model sudah mencapai konvergensi lebih awal (pada epoch 17). Dengan early stopping, pelatihan yang lebih lama dapat dihindari, mengurangi waktu komputasi.