⚡ Wind-Farm Power Regression 💨

Goal

Predict real-time electrical output from turbine-side sensor data.

Metric

Hit **$R^2 \geq 0.85$** on the *hold-out* test set (used exactly once).

Model	Key hyper-params	CV R^2	Test R^2	Notes
Ridge	α = 1.0	~0.79	0.78	Linear baseline – misses non-linear patterns
SVR (RBF)	C = 10, γ = 0.1, ε = 0.1	**0.87**	**0.86**	Clears the bar; handles sensor interactions
MLP (128, 64)	α = 1e-3, lr = 0.001	0.85	0.84	Near target but a touch of over-fit

Chosen model → SVR (RBF) – highest and most stable score.

Modelling Strategy — Why These Choices?

1. **Start simple.** Ridge/Lasso set a linear baseline and expose the non-linearity gap.
2. **Scale everything.** SVR + MLP are sensitive to feature scale → pipeline kicks off with `StandardScaler`.
3. **GridSearchCV (5-fold, R^2).** Brute-force but reliable: balances bias/variance without peeking at test.
4. **SVR vs MLP.** SVR tunes faster on mid-size data; MLP offers a flexible non-linear backup.

5. **Reproducibility.** All CV splits and network inits use `random_state = 42`.

> No exotic ensembles: SVR already meets spec with fewer moving parts.