Statistics and Data Analysis
Unit 05 – Lecture 02 Notes

Tofik Ali

February 14, 2026

Topic

Filter, wrapper, and embedded feature selection methods (overview).

Learning Outcomes

- Explain filter methods (variance, correlation, mutual information)
- Explain wrapper methods (RFE) at a high level
- Explain embedded methods (lasso, tree importance) at a high level
- Discuss pros/cons of each approach

Detailed Notes

These notes are designed to be read alongside the slides. They expand each slide bullet into plain-language explanations, small worked examples, and common pitfalls. When a formula appears, emphasize (1) what each symbol means, (2) the assumptions needed to use it, and (3) how to interpret the final number in the problem context.

Filter Methods

- Fast scoring without training many models
- Examples: variance threshold, correlation with target
- May miss interactions

Wrapper/Embedded

- Wrapper: search subsets using a model (slow)
- Embedded: selection during training (lasso, trees)

Exercises (with Solutions)

Exercise 1: Low variance

If a feature is almost constant, keep it?

Solution

- Usually no; low variance adds little information.

Exercise 2: Redundant features

Two features have corr=0.99. What might you do?

Solution

- Drop one or use regularization/PCA.

Exercise 3: Wrapper trade-off

Why is RFE slower than filters?

Solution

- It trains many models on many subsets.

Exit Question

When would you prefer a fast filter method over a wrapper method?

Demo (Python)

Run from the lecture folder:

```
python demo/demo.py
```

Output files:

- `images/demo.png`
- `data/results.txt`

References

- Montgomery, D. C., & Runger, G. C. *Applied Statistics and Probability for Engineers*, Wiley.
- Devore, J. L. *Probability and Statistics for Engineering and the Sciences*, Cengage.
- McKinney, W. *Python for Data Analysis*, O'Reilly.