The difference comes from **what each method measures**:

* **Spearman correlation** → measures monotonic relationships (whether one variable increases/decreases consistently with another). It doesn’t care about how much info the feature gives beyond monotonicity, and it can be inflated by non-linear but ordered patterns.
* **Mutual Information (MI)** → measures any dependency (linear or non-linear) by looking at shared information, not just ordering. MI can detect relationships that aren’t monotonic — but it’s also sensitive to noise, and very small values mean “almost no predictive power.”

Why you’re seeing a mismatch:

* Spearman can give a non-zero correlation for variables that **don’t really help prediction**, because they might have a weak monotonic trend but still overlap in what they explain.
* MI is stricter in practice — it’s asking “how much uncertainty about the target disappears when I know this feature?”

**Trust level**:

* If your goal is **prediction**, MI is usually more realistic for feature selection.
* If your goal is **understanding monotonic trends**, Spearman is fine — but it may overestimate usefulness.

I’d recommend: **Use MI + domain knowledge** for model building, and use Spearman mainly for understanding feature-target relationships.