CV R² Threshold Analysis: When Good Scores Mean Trouble