

Public: 3rd -> Private: 22nd | Trust Your CV

As the dust settles on the competition, our final results—a Public LB score of 0.38428 dropping to a Private LB score of 0.38387—offer a clear and valuable lesson. This post will detail our final ensemble strategy, but more importantly, it will reflect on how our journey validates the crucial advice famously shared by fellow competitor [Ravi R](#): one must always be prudent and trust their own CV score.

Our Winning Strategy in a Nutshell

Our top solution was not a simple blend but a structured, 10-model hierarchical ensemble. We manually grouped high-performing public submissions into three distinct clusters:

- An "**Elite**" group with the top-scoring models.
- A "**Good**" group of strong, complementary models.
- A single "**Independent**" model, chosen for its diversity to cover the others' blind spots.

The final prediction came from a carefully tuned, weighted vote between representatives from these three groups. This structure proved to be more effective than any other we tested.

The Optimization Journey

This final model was the result of dozens of systematic experiments. We started with simple blends, evolved to the hierarchical structure described above, and then began a long "hill-climbing" process. We meticulously tuned parameters (rank_points, weight_power) and made single, isolated model swaps to find the optimal 10-player roster. This journey taught us that ensemble chemistry is incredibly sensitive, and simply adding a higher-scoring model doesn't always guarantee a better result.

The Final Lesson: "Trust Your CV"

Our final scores—a Public LB of 0.38428 and a Private LB of 0.38387—are a perfect illustration of the advice shared by fellow competitor **Ravi R** during the competition. He wisely cautioned against chasing public LB scores with blends that are not backed by a solid cross-validation (CV) strategy.

Since we were ensembling public submissions, the Public LB was our only form of validation. Our entire optimization process, as methodical as it was, was essentially fine-tuning our submission for that specific public dataset. The small drop in our private score is a classic "shake-down" and the price paid for not having an independent CV score. Our experience perfectly validates Ravi's point: the public LB is a guide, not the ultimate truth.

Thank you to the community for the public work that made this possible. Happy Kagglng!