The paper is interesting. It introduces models in the Netflix Prize. The contest "concerns recommender system for movies" and asked participants to build models predicting the hold-out set of 3 million ratings with the ratings of almost 500,000 people on 18,000 movies (Bell, Koren, & Volinsky, 2010, p.24). Good models should make good use of as much data as possible where data are sufficient while avoid overfitting where data are scarce (Bell et al., 2010, p.25). The criterion is to minimize root mean square error (RMSE), which compared the prediction to the true grade (Bell et al., 2010, p.24). The function is:

$$\text{RMSE(prediction)} = \sqrt{E((prediction - grade)^2)} = \sqrt{\frac{\sum_{t=1}^{T}(prediction_t - grade_t)^2}{T}}$$

The cutoff is 10% improvement (Bell et al., 2010, p.29).

At the beginning of the contest, the most commonly used method was nearest neighbors (Bell et al., 2010, p.25). Intuitively, the rating of similar movies should be highly correlated to each other and thus the best prediction should be a weighted average of similar movies by the same user (since we should control the individual preference) (Bell et al., 2010, p.25). "Similarity is measured via Pearson correlation, cosine similarity, or other metric calculated on the ratings" (Bell et al., 2010, p.25). However, the method had some weaknesses. First, how to decide the similarity is a problem to solve and there is no formal justification for any criterion; second, relative weights don't take compositions of a neighborhood into consideration, leading to the "double counting" problem; finally, it's hard to rate a movie with few or no neighbors with that method (Bell et al., 2010, p.26).

According to the paper, the best predictive models were hybrids of multiple

models (Bell et al., 2010, p.29). A blend will almost certainly improve the prediction, especially when the new components are not highly correlated to the others. So the latter models always tried to combine a lot of prediction sets. The authors also did that but using non-linear combinations to achieve a 10% improvement and win the competition (Bell et al., 2010, p.29).

Reference

Bell, Robert M., Yehuda Koren, and Chris Volinsky, "All Together Now: A Perspective on the Netflix Prize", *Chance*, 2010, 23 (1), 24-29.