What is the difference:

**Random Forest** is a **bagging** algorithm rather than a **boosting** algorithm. They are two opposite ways to achieve a low error.

1). **Bagging**

Is a powerful method to improve the performance of simple models and reduce the overfitting of more complex models. The principle is straightforward to understand; instead of fitting the model on one population sample, several models are done on different samples (with replacement) of the population.

Then, these models are aggregated by using their average, weighted average, or voting system mainly for **Classification** (Votes are given to each prediction, and the prediction gets maximum votes to get selected) vs. **Regression** (The average value is used to predict the statistics).

2). **Random Forest**

reduces the variance of many "complex" models with low bias. The fundamental difference between bagging and Random Forest is that in Random Forest, only a subset of features are selected at random out of the total, and the best split feature from the subset is used to spit each mode in a tree, unlike bagging, where all parts are considered for splitting a node.

3). **Boosting**

we apply training models or learners to the training set iteratively, and weights are assigned to each class according to probability.

Initially, equal probabilities are assigned to each class. After 1st iteration, the misclassified items are assigned more probability than the correctly classified items. The process is repeated.

Finally, we will combine the outputs of different learners through voting. While voting, we assign weights to each learner depending on how good the learning model was.

See: <https://colab.research.google.com/drive/1gyizPhqN5NB0ZaWNBB8uH9BAGytWUcOD>