

هي خوارزمية للتعلم الآلي تم تطويرها بناءً على مجموعة من أشجار القرار. تستخدم هذه الخوارزمية للتصنيف، عادةً ما تتمتع خوارزمية الغابة العشوائية بدقة أفضل مقارنة بشجرة القرار. حيث تنبثق كل شجرة فردية في الغابة العشوائية من توقع الصنف ويصبح الصنف الذي يحصل على أكبر عدد من الأصوات هو توقع النموذج. Random Forest السبب في نجاح هذه خوارزمية في علم البيانات يرجع الى سبب أن العدد الكبير من النماذج (الأشجار) غير المترابطة نسبياً التي تعمل كلجنة سيتتفوق على أي نموذج مكون من شجرة قرار واحدة. معامل الارتباط المنخفض بين نماذج أشجار القرار هو المفتاح. تماماً مثل الاستثمارات ذات معامل الارتباط المنخفض (مثل الأسهم والسندات من شركات مختلفة المجالات) التي تشكل المحفظة الإستثمارية. يمكن للنماذج غير المترابطة أن تنتج تنبؤات مجمعة أكثر دقة من أي تنبؤات فردية. سبب هذا التأثير الرائع هو أن الأشجار تحمي بعضها البعض من أخطائها الفردية (طالما أنها لا تخطئ جميعاً في نفس الاتجاه باستمرار). في حين أن بعض الأشجار قد تكون خاطئة، إلا أن العديد من الأشجار الأخرى ستكون صحيحة، بحيث يمكن للأشجار كمجموعة أن تتحرك في الاتجاه الصحيح.