

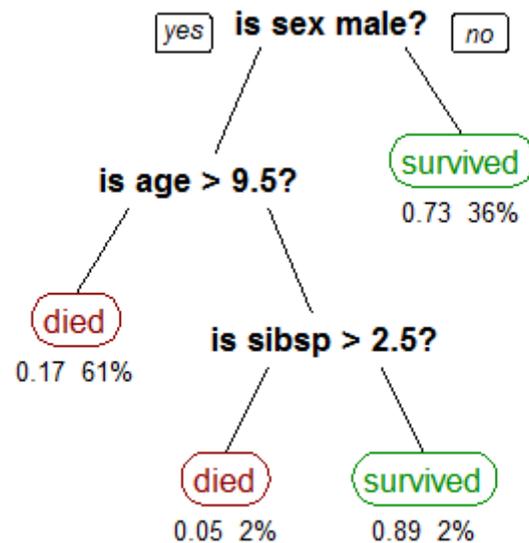
# Случайный лес (Random forest)

Ставровская Елена

2016

# Дерево принятия решений

- Узел – один из признаков
- Ребро – значение признака
- Лист – значение целевой функции(класс)



# Дерево принятия решений

Построение:

- Выбираем очередной признак  $P$ , помещаем его в корень.
- Для всех его значений  $v$ :
  - Оставляем из тестовых примеров только те, у которых значение атрибута  $P$  равно  $v$
  - Рекурсивно строим дерево в этом потомке

# Дерево принятия решений

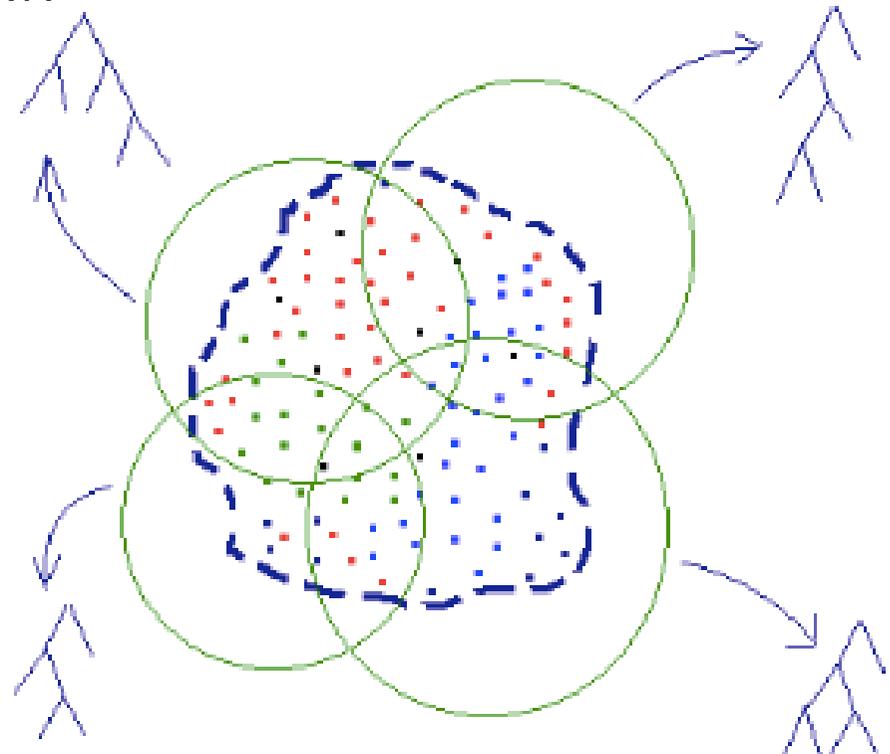
Проблема:

Жадный алгоритм, результат зависит от  
порядка выбора признаков



# Случайный лес (Random Forest)

- Создаем лес случайных решающих деревьев:
  - Выбираем из обучающей выборки  $N$  объектов с **повторением**
  - Строим дерево решений



# Случайный лес (Random Forest)

- Классификация объектов проводится путём голосования: каждое дерево леса относит классифицируемый объект к одному из классов, и побеждает класс, за который проголосовало наибольшее число деревьев.



# Random Forest в R

```
library(randomForest)
```

```
model = randomForest(Type ~ ., train)
```

```
#важность переменных для классификатора
```

```
importance(model)
```

```
#вытащим одно из деревьев леса
```

```
getTree(model, k=10, labelVar=T)
```