Variability-Aware Performance Prediction: A Statistical Learning Approach

Jianmei Guo
Krzysztof Czarnecki

University of Waterloo
Canada

Sven Apel
Norbert Siegmund

University of Passau
Germany

Andrzej Wąsowski

IT University of Copenhagen
Denmark

Configurable software and variability are ubiquitous

Configure software to tailor functional behavior

```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

Configure software to tailor functional behavior

a command-line tool to encode a video stream

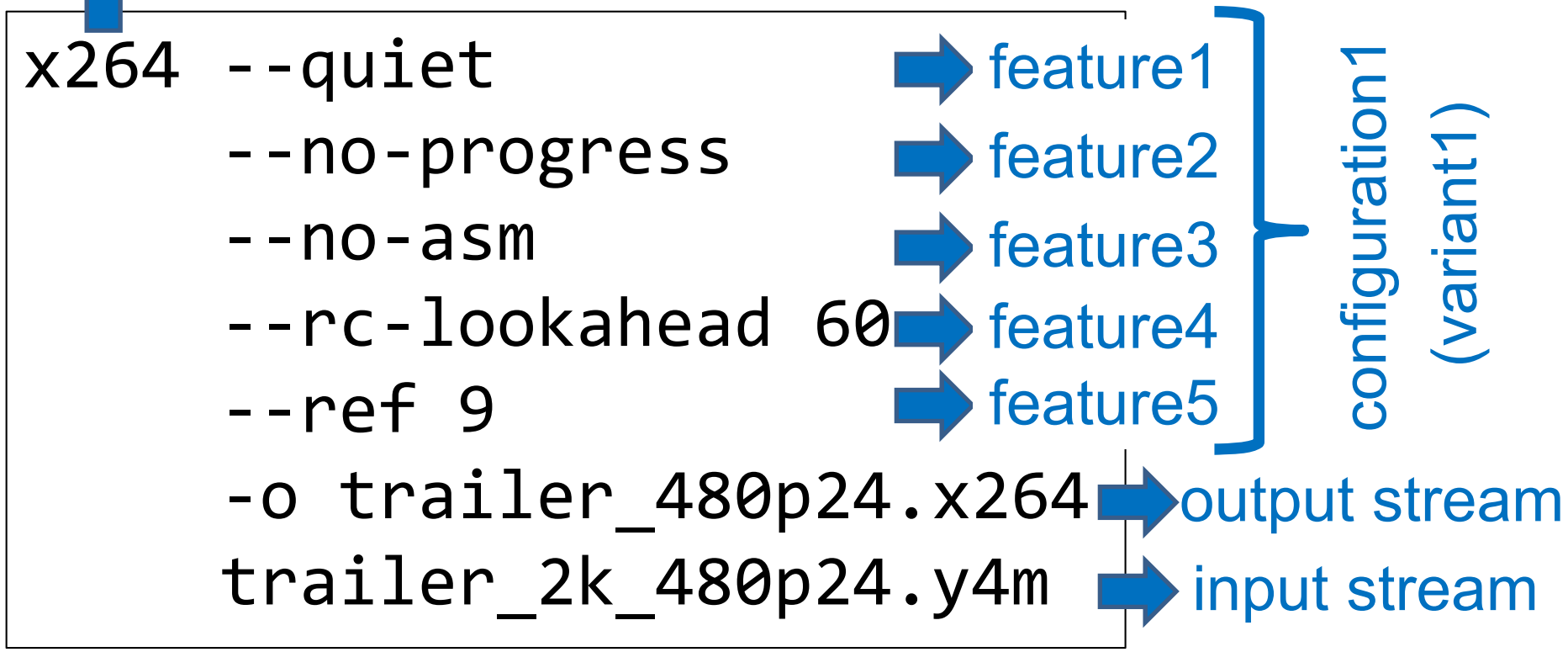
```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

output stream

input stream

Configure software to tailor functional behavior

a command-line tool to encode a video stream



Configure software to meet a certain performance goal

configuration1

```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

424 seconds

configuration2

```
x264
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

651 seconds

Measure the performance of all configurations?

- An exponential number of configurations
  - N features → about 2^N configurations
- The cost of measurement may be high

Feature-wise measurement?

configuration1

```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

324 seconds

configuration2

```
x264
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

551 seconds

P(**quiet**)

= 551 – 324

= **227** seconds

Feature-wise measurement?

configuration1

```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

324 seconds

configuration2

```
x264
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

551 seconds

P(**quiet**)
= 551 – 324
= **227** seconds

configuration3

```
x264 --quiet
     --no-progress

     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

487 seconds

configuration4

```
x264
     --no-progress

     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

661 seconds

P'(**quiet**)
= 661– 487
= **174** seconds

Feature-wise measurement?

configuration1

```
x264 --quiet
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

324 seconds

configuration2

```
x264
     --no-progress
     --no-asm
     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

551 seconds

configuration3

```
x264 --quiet
     --no-progress

     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

487 seconds

configuration4

```
x264
     --no-progress

     --rc-lookahead 60
     --ref 9
     -o trailer_480p24.x264
     trailer_2k_480p24.y4m
```

661 seconds

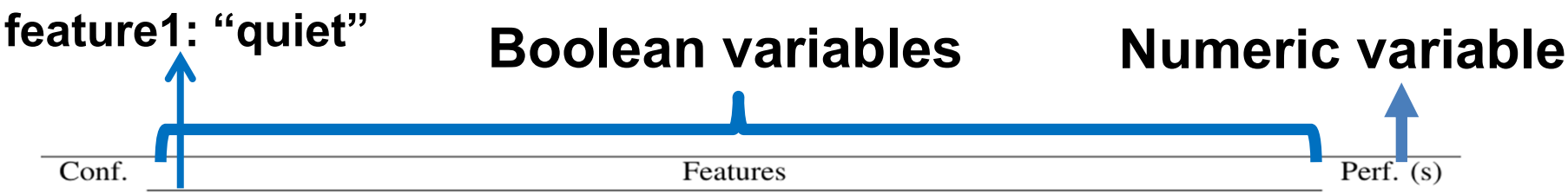
P(**quiet**)
= 551 – 324
= **227** seconds

P'(**quiet**)
= 661– 487
= **174** seconds

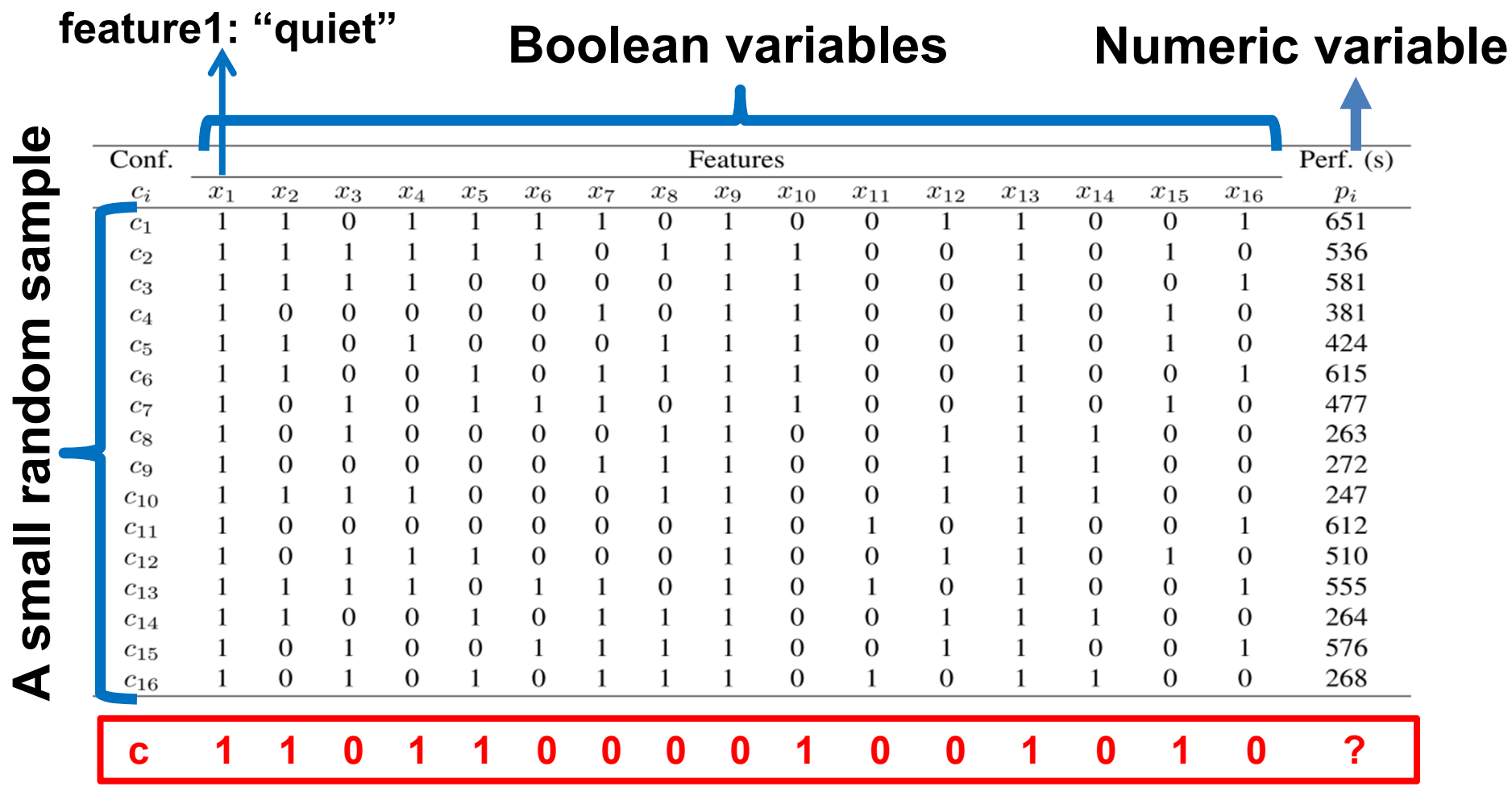
[Siegmund et al, ICSE'12]

Is it feasible to use
a **small random sample** of measured configurations
as a basis to **accurately predict**
the **performance** of other configurations,
especially when **features interact**?

Conf.	Features																Perf. (s)
c_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	p_i
c_1	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	1	651
c_2	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	536
c_3	1	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	581
c_4	1	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	381
c_5	1	1	0	1	0	0	0	1	1	1	0	0	1	0	1	0	424
c_6	1	1	0	0	1	0	1	1	1	1	0	0	1	0	0	1	615
c_7	1	0	1	0	1	1	1	0	1	1	0	0	1	0	1	0	477
c_8	1	0	1	0	0	0	0	1	1	0	0	1	1	1	0	0	263
c_9	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	272
c_{10}	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	0	247
c_{11}	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	612
c_{12}	1	0	1	1	1	0	0	0	1	0	0	1	1	0	1	0	510
c_{13}	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	555
c_{14}	1	1	0	0	1	0	1	1	1	0	0	1	1	1	0	0	264
c_{15}	1	0	1	0	0	1	1	1	1	0	0	1	1	0	0	1	576
c_{16}	1	0	1	0	1	0	1	1	1	0	1	0	1	1	0	0	268



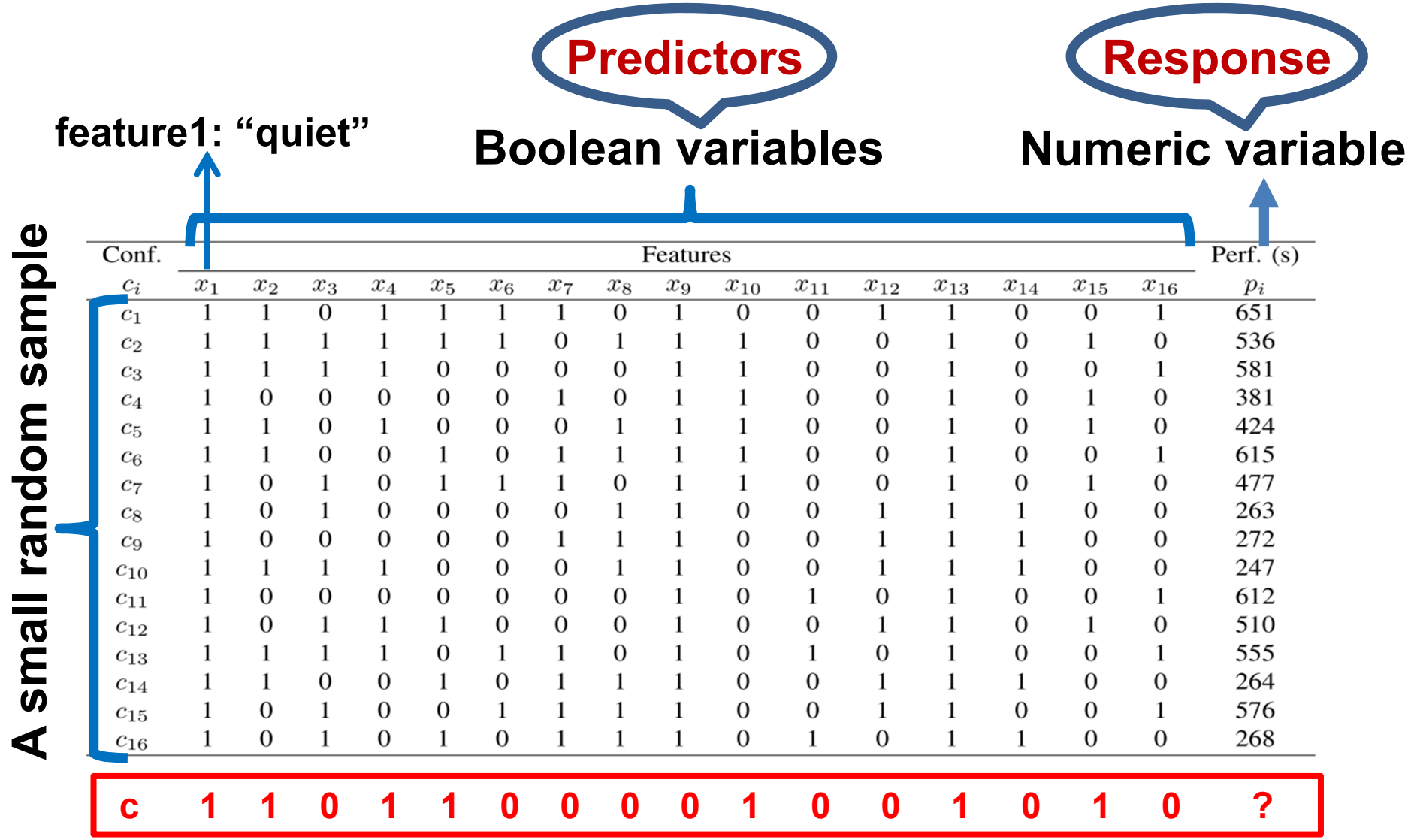
Conf.	Features																Perf. (s)
c_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	p_i
c_1	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	1	651
c_2	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	536
c_3	1	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	581
c_4	1	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	381
c_5	1	1	0	1	0	0	0	1	1	1	0	0	1	0	1	0	424
c_6	1	1	0	0	1	0	1	1	1	1	0	0	1	0	0	1	615
c_7	1	0	1	0	1	1	1	0	1	1	0	0	1	0	1	0	477
c_8	1	0	1	0	0	0	0	1	1	0	0	1	1	1	0	0	263
c_9	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	272
c_{10}	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	0	247
c_{11}	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	612
c_{12}	1	0	1	1	1	0	0	0	1	0	0	1	1	0	1	0	510
c_{13}	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	555
c_{14}	1	1	0	0	1	0	1	1	1	0	0	1	1	1	0	0	264
c_{15}	1	0	1	0	0	1	1	1	1	0	0	1	1	0	0	1	576
c_{16}	1	0	1	0	1	0	1	1	1	0	1	0	1	1	0	0	268



Conf.	Features																Perf. (s)
c_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	p_i
c_1	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	1	651
c_2	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	536
c_3	1	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	581
c_4	1	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	381
c_5	1	1	0	1	0	0	0	1	1	1	0	0	1	0	1	0	424
c_6	1	1	0	0	1	0	1	1	1	1	0	0	1	0	0	1	615
c_7	1	0	1	0	1	1	1	0	1	1	0	0	1	0	1	0	477
c_8	1	0	1	0	0	0	0	1	1	0	0	1	1	1	0	0	263
c_9	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	272
c_{10}	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	0	247
c_{11}	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	612
c_{12}	1	0	1	1	1	0	0	0	1	0	0	1	1	0	1	0	510
c_{13}	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	555
c_{14}	1	1	0	0	1	0	1	1	1	0	0	1	1	1	0	0	264
c_{15}	1	0	1	0	0	1	1	1	1	0	0	1	1	0	0	1	576
c_{16}	1	0	1	0	1	0	1	1	1	0	1	0	1	1	0	0	268

c	1	1	0	1	1	0	0	0	0	1	0	0	1	0	1	0	?

A non-linear regression problem

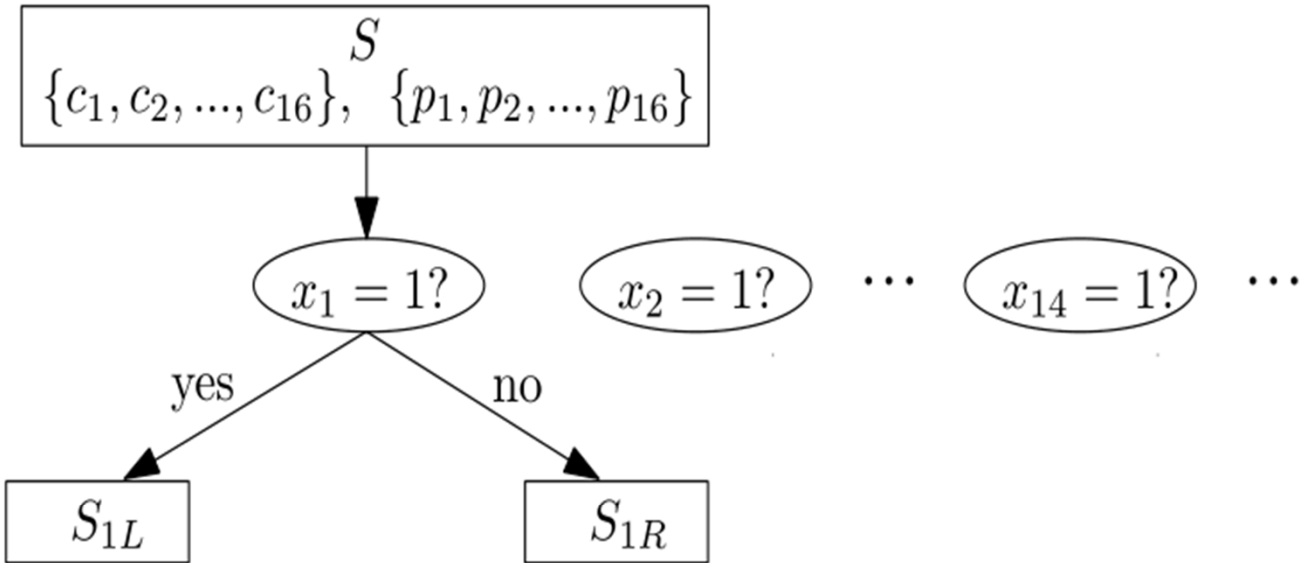


Conf.	Features																Perf. (s)
c_i	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	p_i
c_1	1	1	0	1	1	1	1	0	1	0	0	1	1	0	0	1	651
c_2	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	536
c_3	1	1	1	1	0	0	0	0	1	1	0	0	1	0	0	1	581
c_4	1	0	0	0	0	0	1	0	1	1	0	0	1	0	1	0	381
c_5	1	1	0	1	0	0	0	1	1	1	0	0	1	0	1	0	424
c_6	1	1	0	0	1	0	1	1	1	1	0	0	1	0	0	1	615
c_7	1	0	1	0	1	1	1	0	1	1	0	0	1	0	1	0	477
c_8	1	0	1	0	0	0	0	1	1	0	0	1	1	1	0	0	263
c_9	1	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	272
c_{10}	1	1	1	1	0	0	0	1	1	0	0	1	1	1	0	0	247
c_{11}	1	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	612
c_{12}	1	0	1	1	1	0	0	0	1	0	0	1	1	0	1	0	510
c_{13}	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	1	555
c_{14}	1	1	0	0	1	0	1	1	1	0	0	1	1	1	0	0	264
c_{15}	1	0	1	0	0	1	1	1	1	0	0	1	1	0	0	1	576
c_{16}	1	0	1	0	1	0	1	1	1	0	1	0	1	1	0	0	268

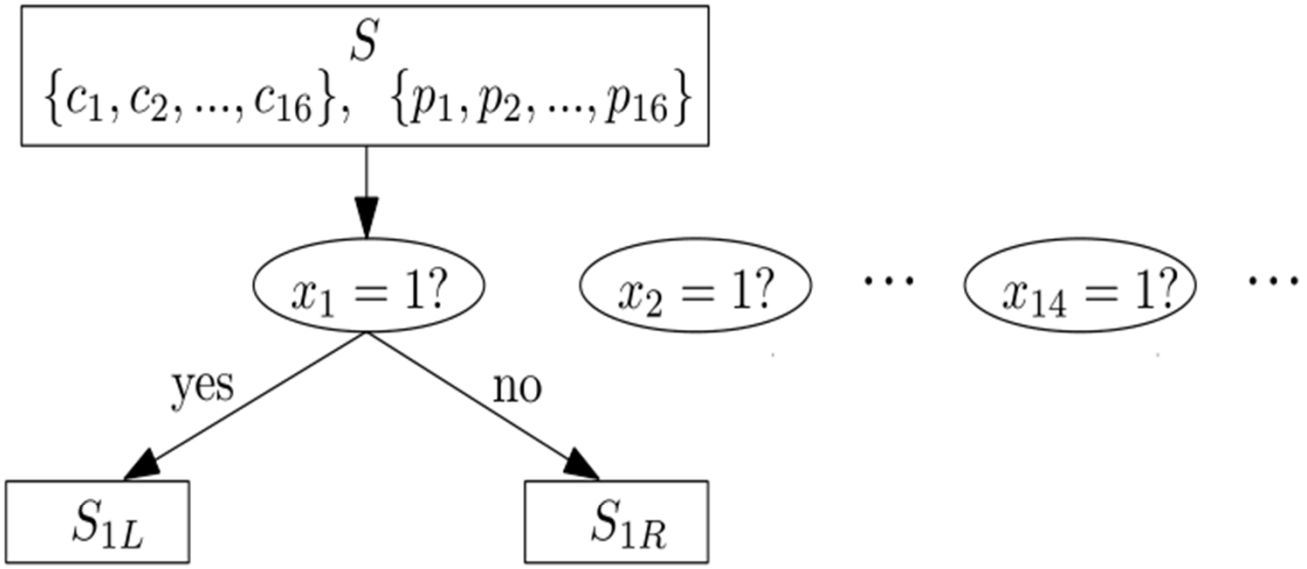
c	1	1	0	1	1	0	0	0	0	1	0	0	1	0	1	0	?

Classification And Regression Trees (CART) **[Breiman et al, 1984]**

Classification And Regression Trees (CART) **[Breiman et al, 1984]**

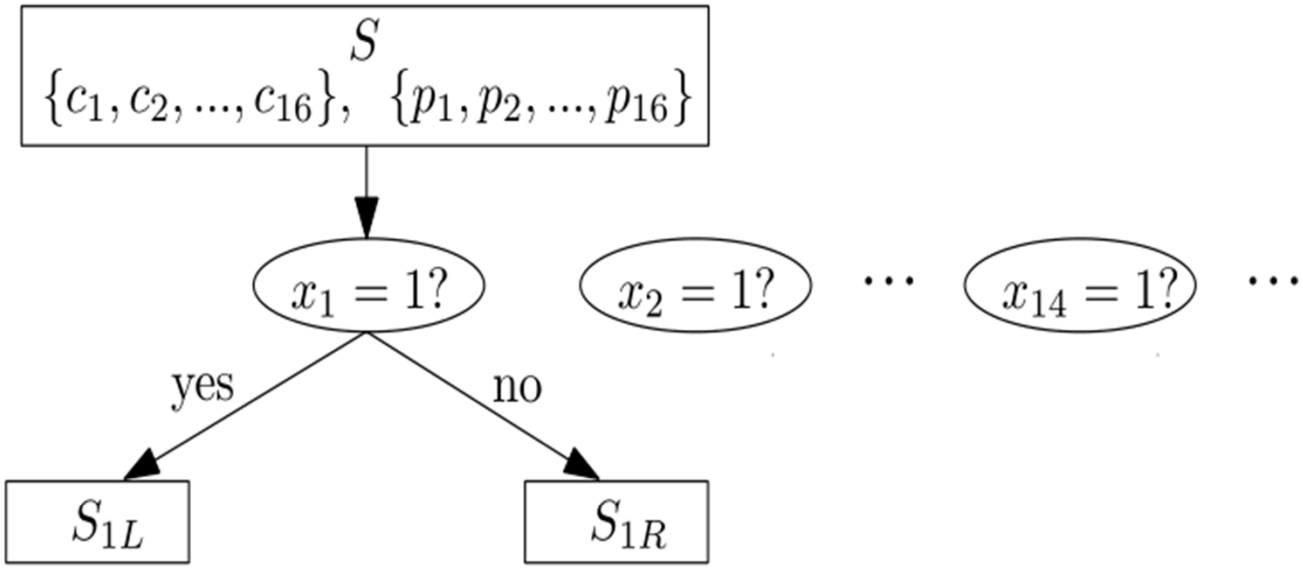


Classification And Regression Trees (CART) [Breiman et al, 1984]



Sample mean $\ell_{S_{iL}} = \frac{1}{|S_{iL}|} \sum_{p_j \in S_{iL}} p_j$ $\ell_{S_{iR}} = \frac{1}{|S_{iR}|} \sum_{p_j \in S_{iR}} p_j$

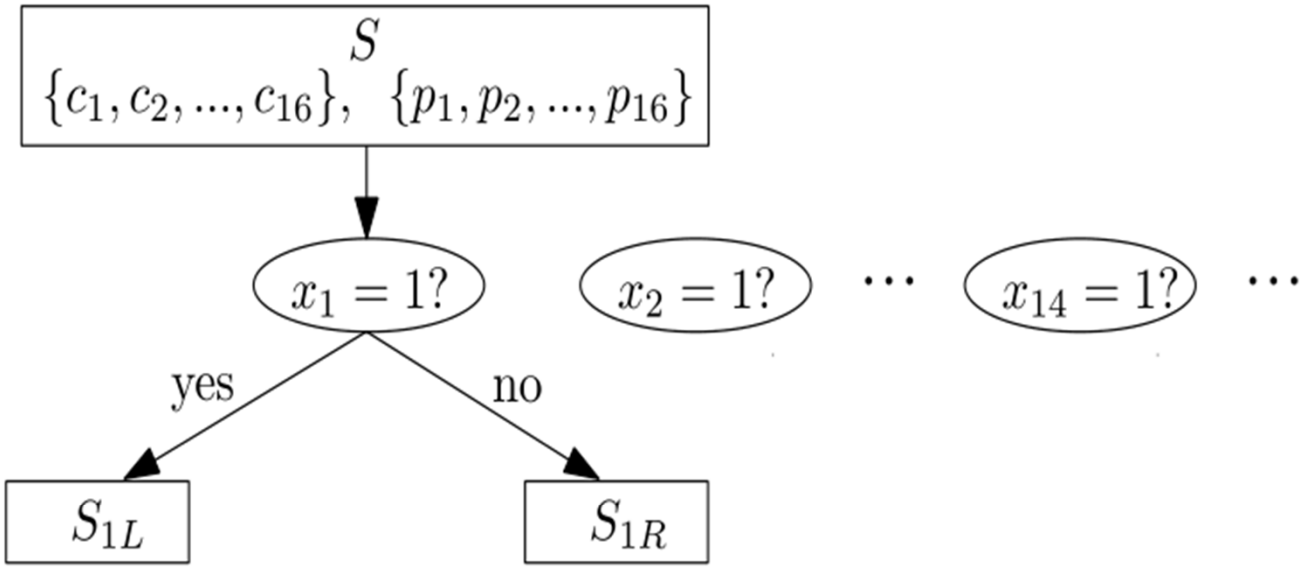
Classification And Regression Trees (CART) [Breiman et al, 1984]



Sample mean $\ell_{S_{iL}} = \frac{1}{|S_{iL}|} \sum_{p_j \in S_{iL}} p_j$ $\ell_{S_{iR}} = \frac{1}{|S_{iR}|} \sum_{p_j \in S_{iR}} p_j$

Squared error loss $\sum_{p_j \in S_{iL}} (p_j - \ell_{S_{iL}})^2$ $\sum_{p_j \in S_{iR}} (p_j - \ell_{S_{iR}})^2$

Classification And Regression Trees (CART) [Breiman et al, 1984]

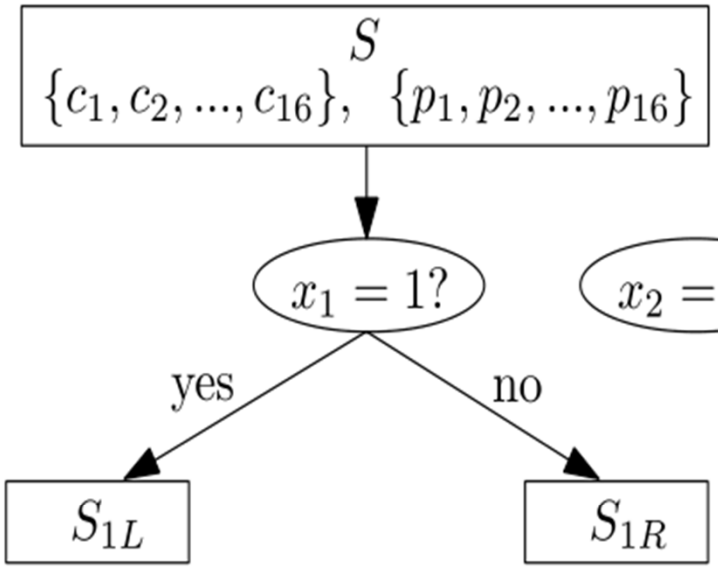


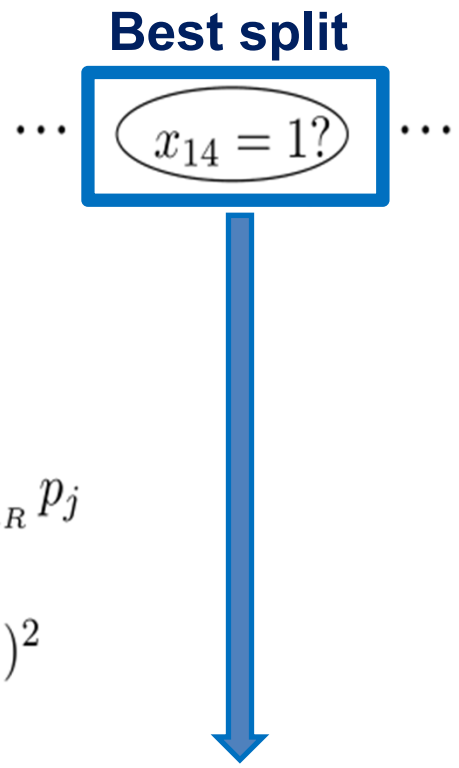
Sample mean $\ell_{S_{iL}} = \frac{1}{|S_{iL}|} \sum_{p_j \in S_{iL}} p_j$ $\ell_{S_{iR}} = \frac{1}{|S_{iR}|} \sum_{p_j \in S_{iR}} p_j$

Squared error loss $\sum_{p_j \in S_{iL}} (p_j - \ell_{S_{iL}})^2$ $\sum_{p_j \in S_{iR}} (p_j - \ell_{S_{iR}})^2$

Sum of squared error loss $\sum_{p_j \in S_{iL}} (p_j - \ell_{S_{iL}})^2 + \sum_{p_j \in S_{iR}} (p_j - \ell_{S_{iR}})^2$ **is minimal**

Classification And Regression Trees (CART) [Breiman et al, 1984]

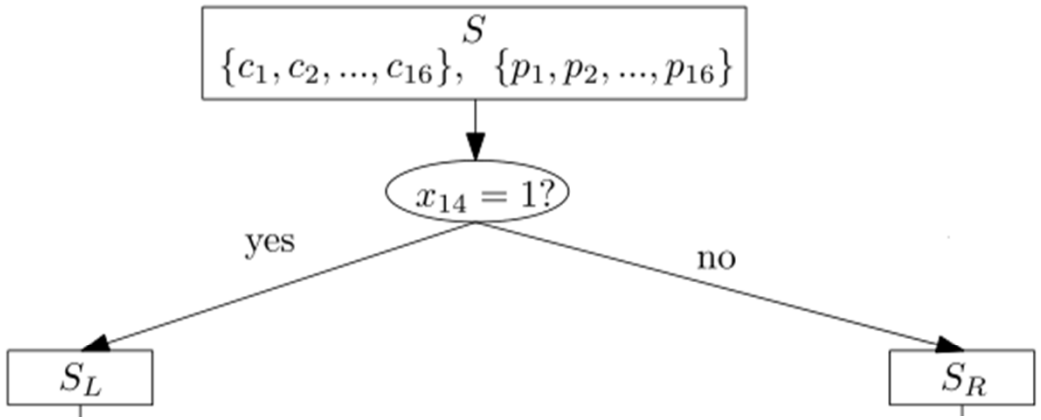


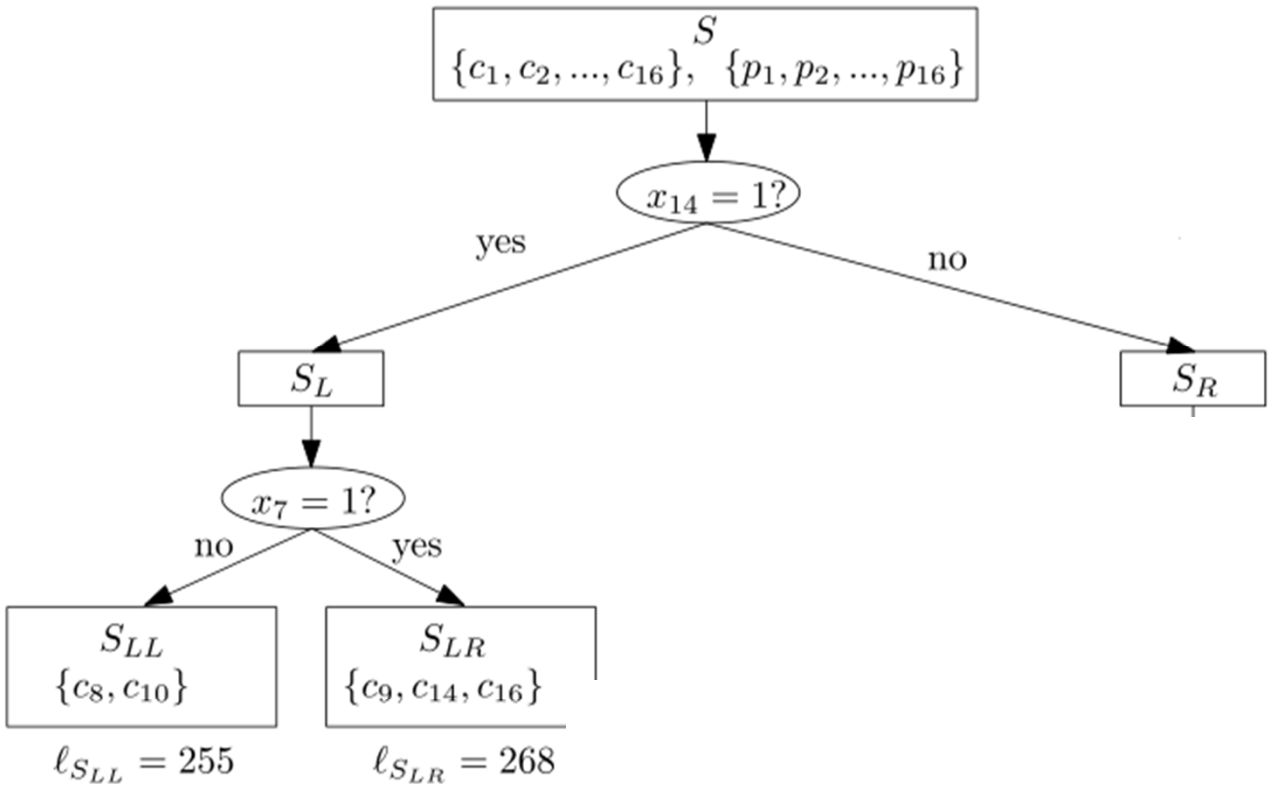


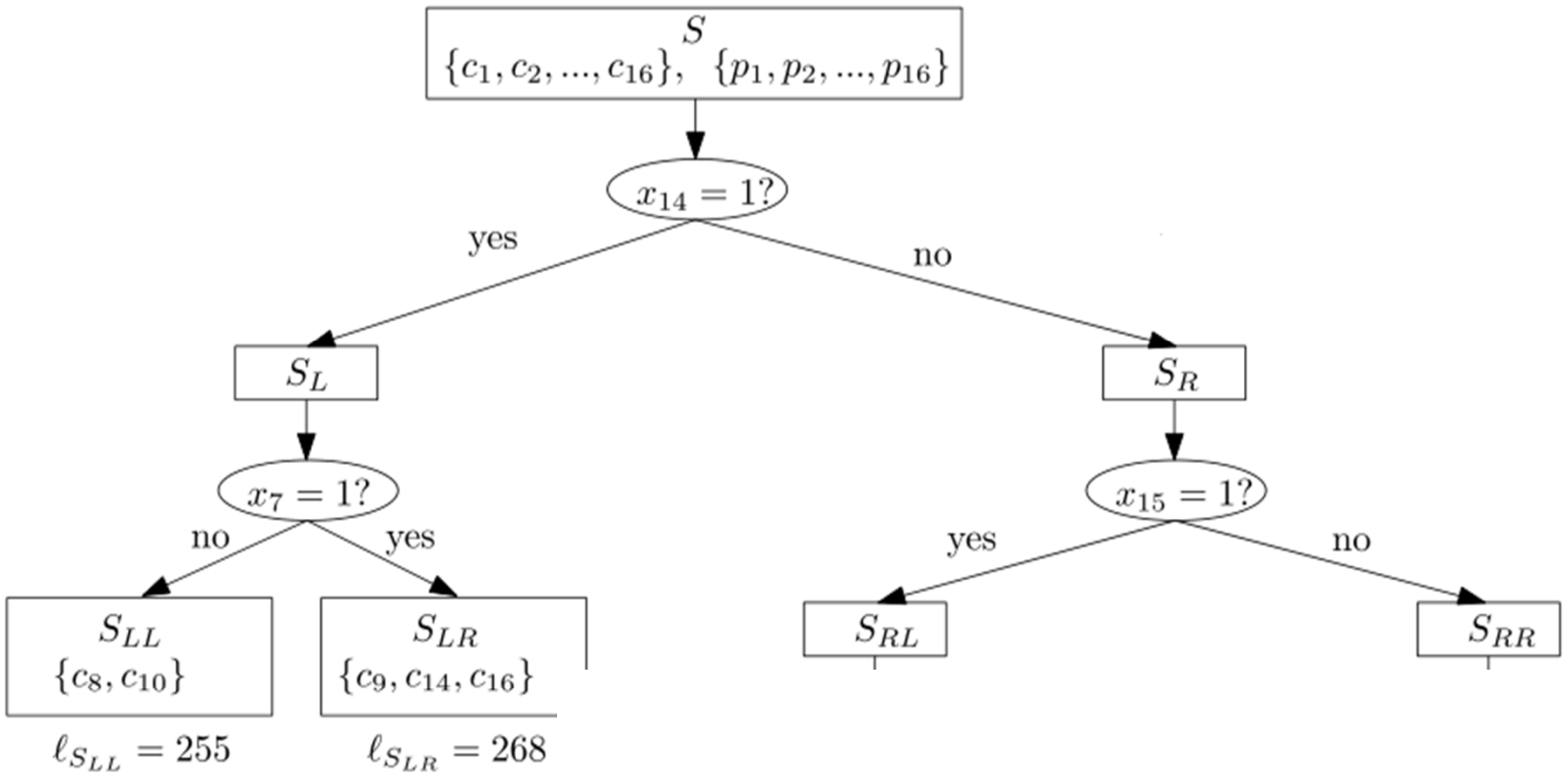
Sample mean $\ell_{S_{iL}} = \frac{1}{|S_{iL}|} \sum_{p_j \in S_{iL}} p_j$ $\ell_{S_{iR}} = \frac{1}{|S_{iR}|} \sum_{p_j \in S_{iR}} p_j$

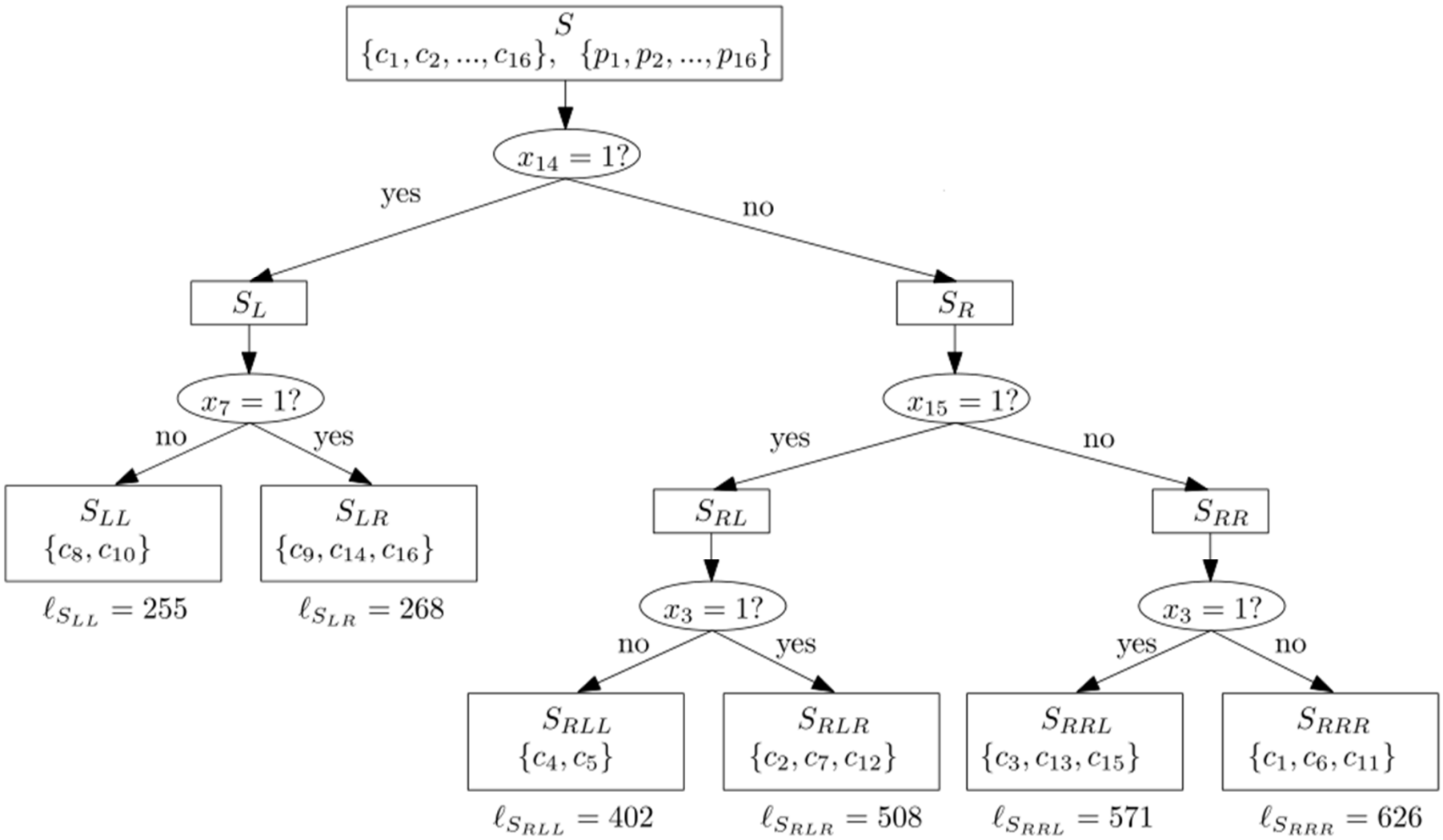
Squared error loss $\sum_{p_j \in S_{iL}} (p_j - \ell_{S_{iL}})^2$ $\sum_{p_j \in S_{iR}} (p_j - \ell_{S_{iR}})^2$

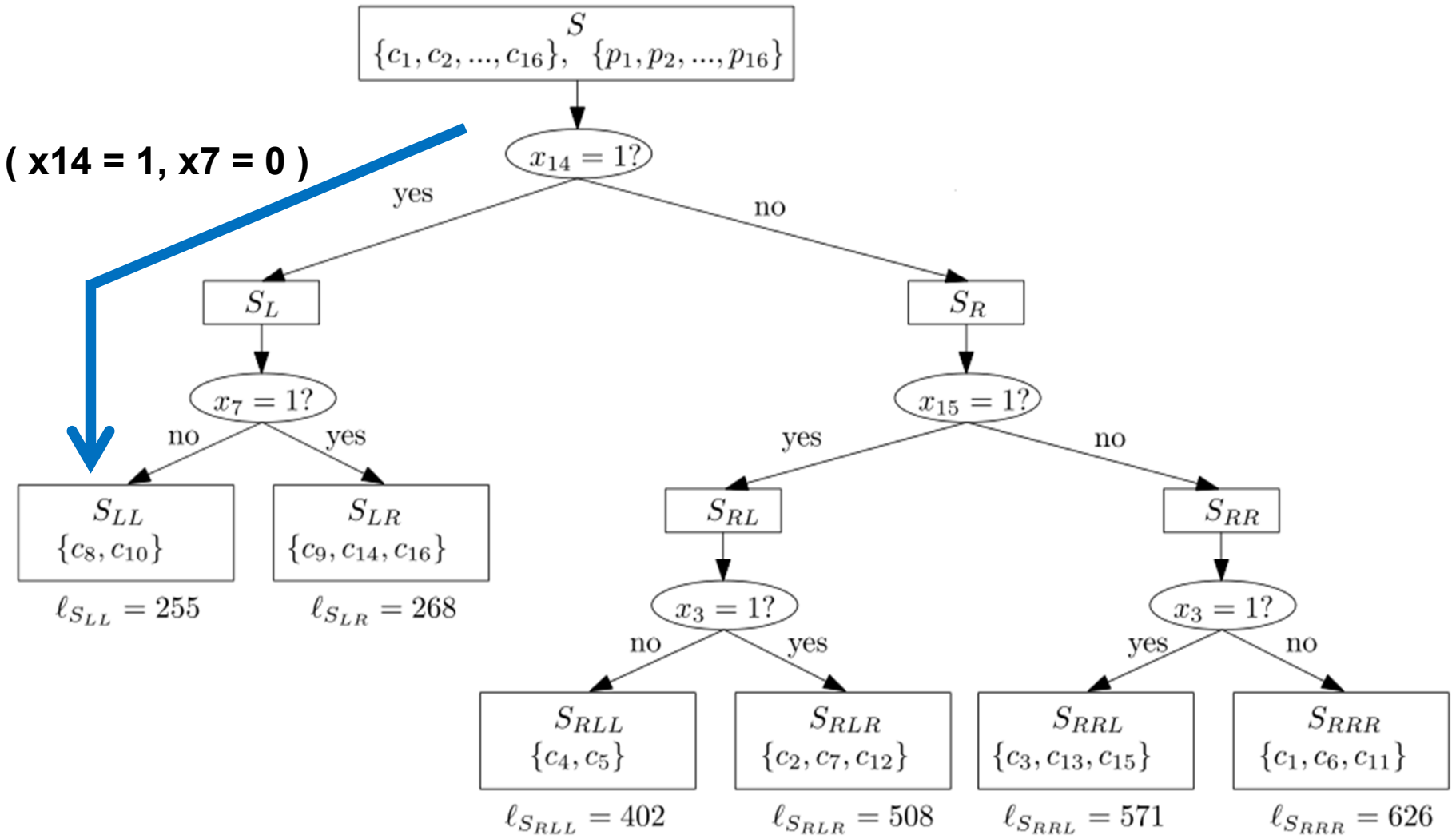
Sum of squared error loss $\sum_{p_j \in S_{iL}} (p_j - \ell_{S_{iL}})^2 + \sum_{p_j \in S_{iR}} (p_j - \ell_{S_{iR}})^2$

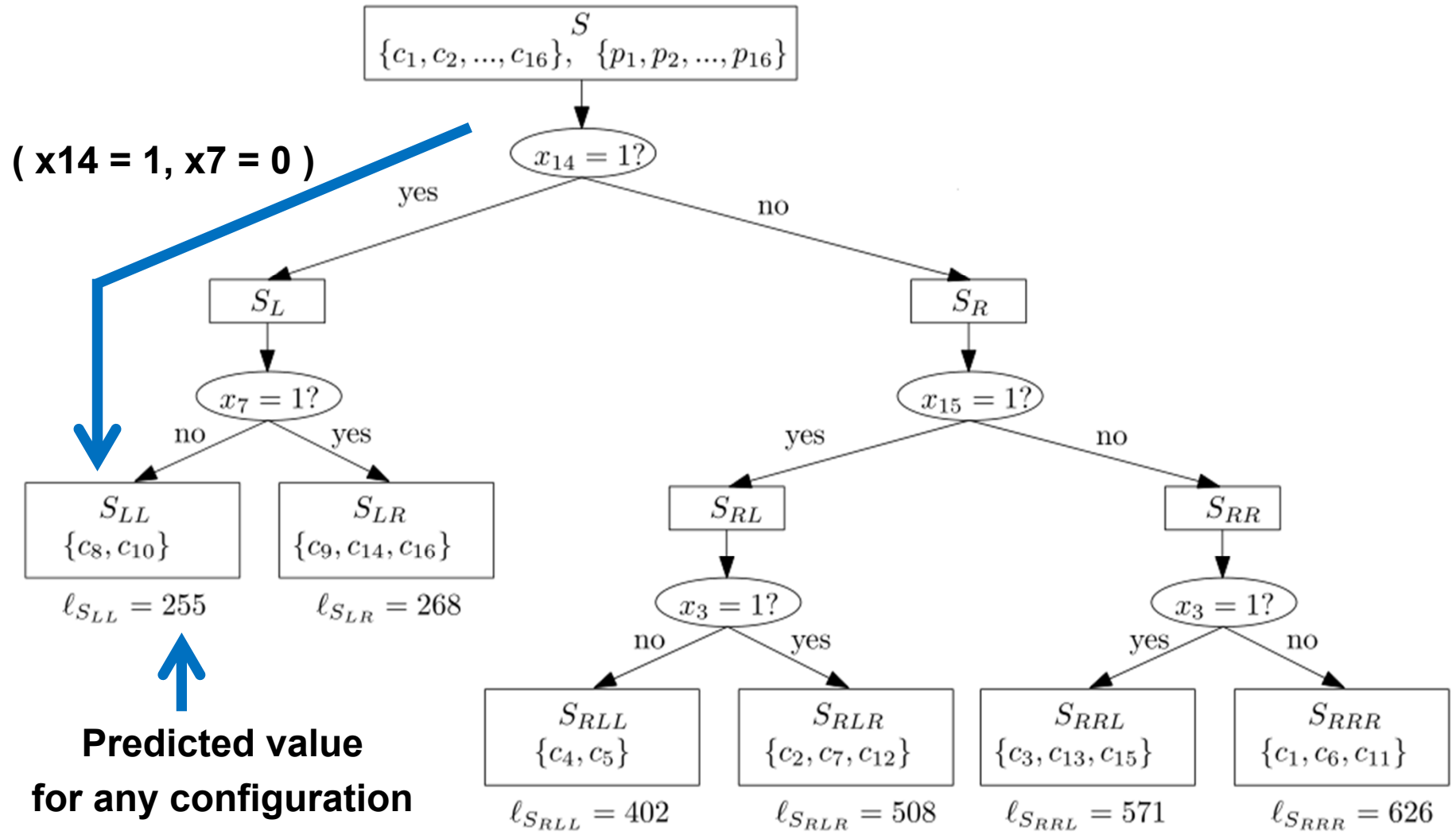

S
$\{c_1, c_2, ..., c_{16}\}, \{p_1, p_2, ..., p_{16}\}$
$x_{14} = 1?$
yes
no
S_L
S_R


S
$\{c_1, c_2, ..., c_{16}\}, \{p_1, p_2, ..., p_{16}\}$
$x_{14} = 1?$
yes
no
S_L
S_R
$x_7 = 1?$
no
yes
S_{LL}
$\{c_8, c_{10}\}$
S_{LR}
$\{c_9, c_{14}, c_{16}\}$
$\ell_{S_{LL}} = 255$
$\ell_{S_{LR}} = 268$


S
$\{c_1, c_2, ..., c_{16}\}, \{p_1, p_2, ..., p_{16}\}$
$x_{14} = 1?$
yes
no
S_L
S_R
$x_7 = 1?$
$x_{15} = 1?$
no
yes
yes
no
S_{LL}
$\{c_8, c_{10}\}$
S_{LR}
$\{c_9, c_{14}, c_{16}\}$
S_{RL}
S_{RR}
$\ell_{S_{LL}} = 255$
$\ell_{S_{LR}} = 268$


S
$\{c_1, c_2, ..., c_{16}\}$, $\{p_1, p_2, ..., p_{16}\}$
$x_{14} = 1?$
yes
no
S_L
S_R
$x_7 = 1?$
$x_{15} = 1?$
no
yes
yes
no
S_{LL}
$\{c_8, c_{10}\}$
S_{LR}
$\{c_9, c_{14}, c_{16}\}$
S_{RL}
S_{RR}
$\ell_{S_{LL}} = 255$
$\ell_{S_{LR}} = 268$
$x_3 = 1?$
$x_3 = 1?$
no
yes
yes
no
S_{RLL}
$\{c_4, c_5\}$
S_{RLR}
$\{c_2, c_7, c_{12}\}$
S_{RRL}
$\{c_3, c_{13}, c_{15}\}$
S_{RRR}
$\{c_1, c_6, c_{11}\}$
$\ell_{S_{RLL}} = 402$
$\ell_{S_{RLR}} = 508$
$\ell_{S_{RRL}} = 571$
$\ell_{S_{RRR}} = 626$


(x14 = 1, x7 = 0)
S
$\{c_1, c_2, ..., c_{16}\}, \{p_1, p_2, ..., p_{16}\}$
$x_{14} = 1?$
yes
no
S_L
S_R
$x_7 = 1?$
$x_{15} = 1?$
no
yes
yes
no
S_{LL}
$\{c_8, c_{10}\}$
S_{LR}
$\{c_9, c_{14}, c_{16}\}$
S_{RL}
S_{RR}
$x_3 = 1?$
$x_3 = 1?$
no
yes
yes
no
S_{RLL}
$\{c_4, c_5\}$
S_{RLR}
$\{c_2, c_7, c_{12}\}$
S_{RRL}
$\{c_3, c_{13}, c_{15}\}$
S_{RRR}
$\{c_1, c_6, c_{11}\}$
$\ell_{S_{LL}} = 255$
$\ell_{S_{LR}} = 268$
$\ell_{S_{RLL}} = 402$
$\ell_{S_{RLR}} = 508$
$\ell_{S_{RRL}} = 571$
$\ell_{S_{RRR}} = 626$



Predicted value for any configuration selecting feature_14 and deselecting feature_7

Evaluation on six real-world configurable software systems

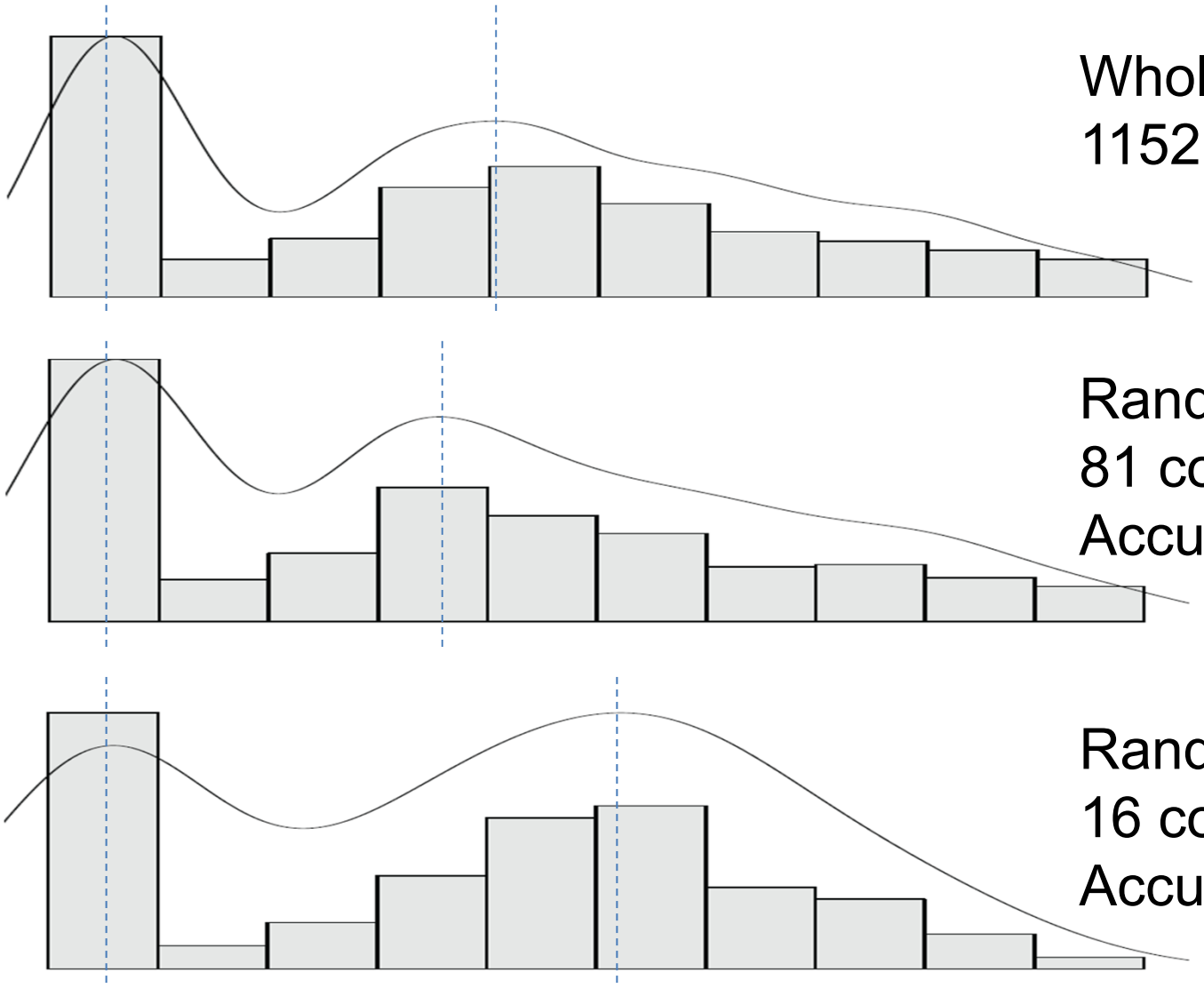
	System	Domain	Language	LOC	#Configurations	#Features
1	APACHE	Web Server	C	230,277	192	9
2	LLVM	Compiler	C++	47,549	1,024	11
3	X264	Encoder	C	45,743	1,152	16
4	BERKELEY DB	Database	C	219,811	2,560	18
5	BERKELEY DB	Database	JAVA	42,596	400	26
6	SQLITE	Database	C	312,625	3,932,160	39

Accuracy and Speed

- Average **94%** accuracy based on **small random** samples
  - sample size is **linear** in the number of features
- **Fully-automated** and **fast**
  - at most **90 milliseconds** in all evaluated case studies

See paper for more details

Why CART works?



Whole population
1152 configurations

Random sample 1
81 configurations
Accuracy: 93.6%

Random sample 2
16 configurations
Accuracy: 84.9%

Comparison to existing methods

	SPL Conqueror [Siegmund et al, ICSE'12]	Our approach
Accuracy	Average 95%	Average 94%
Sample dependency	Specific samples	Random samples
Supporting feature types	Boolean	Boolean and numeric
Running mechanism	Non-progressive	Progressive

Threats to validity

- Generality to other systems
- Empirical parameter tuning for stopping recursive partitioning
- Quantitative study on the similarity between a sample and the whole population

Future work

- More software systems
- Systematic parameter tuning
- Similarity metrics between a sample and the whole population
  - Combining feature selections and performance values
- Other regression techniques

Conclusion

A **variability-aware** approach to performance prediction via a **well-established** statistical learning technique

- works **automatically** and **progressively** with **random** samples
- considers **all** features and identifies the **performance-relevant** ones
- treats selected and deselected features **equally**
- can be **easily** understood, implemented and deployed, **without** feature interaction detection

Thank you!

http://gsd.uwaterloo.ca/cpm