

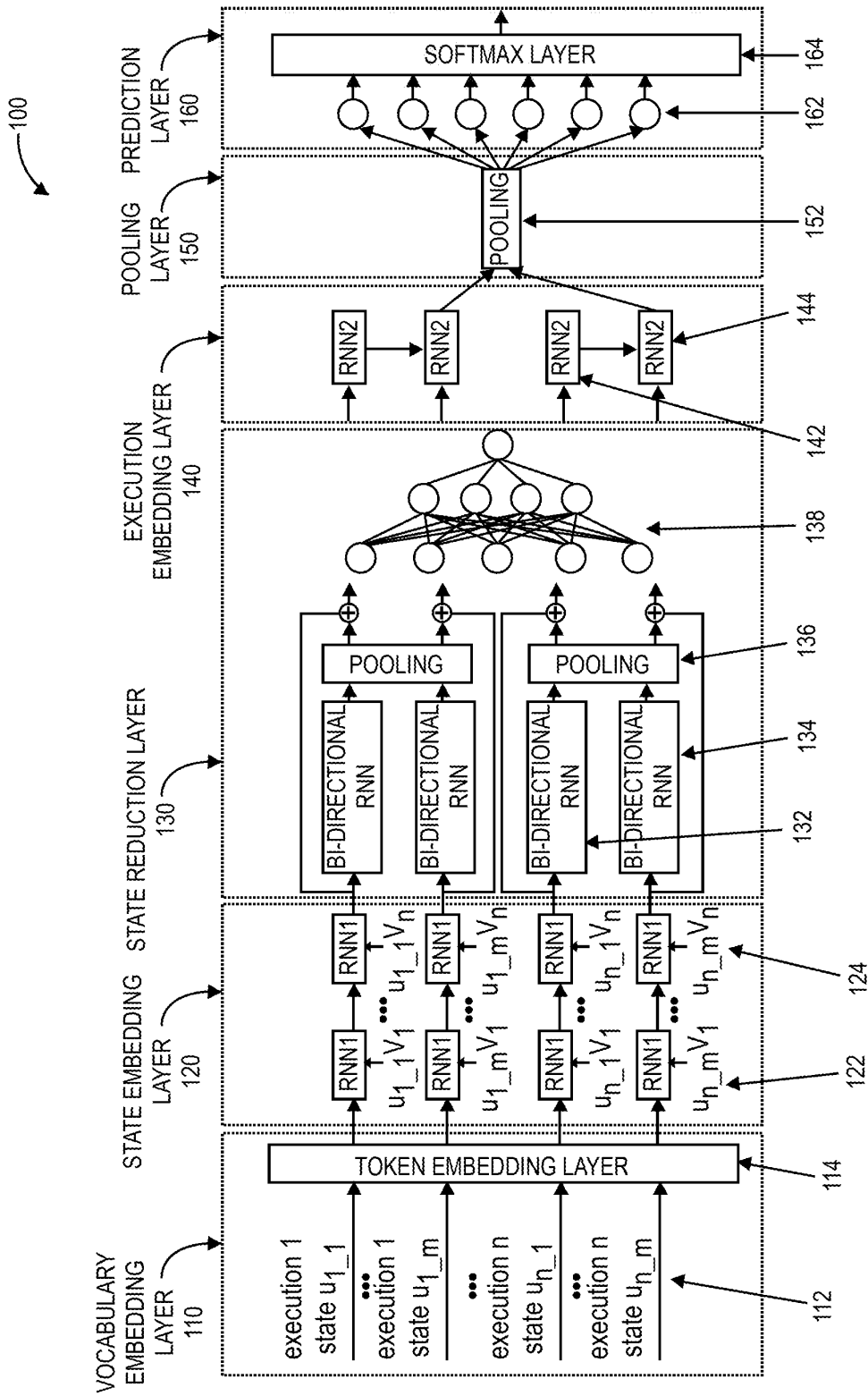
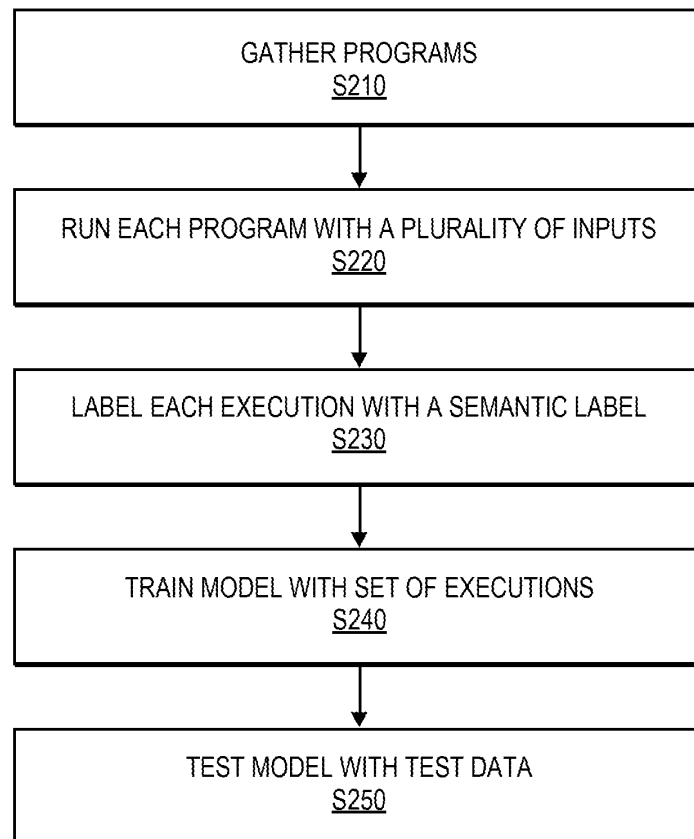
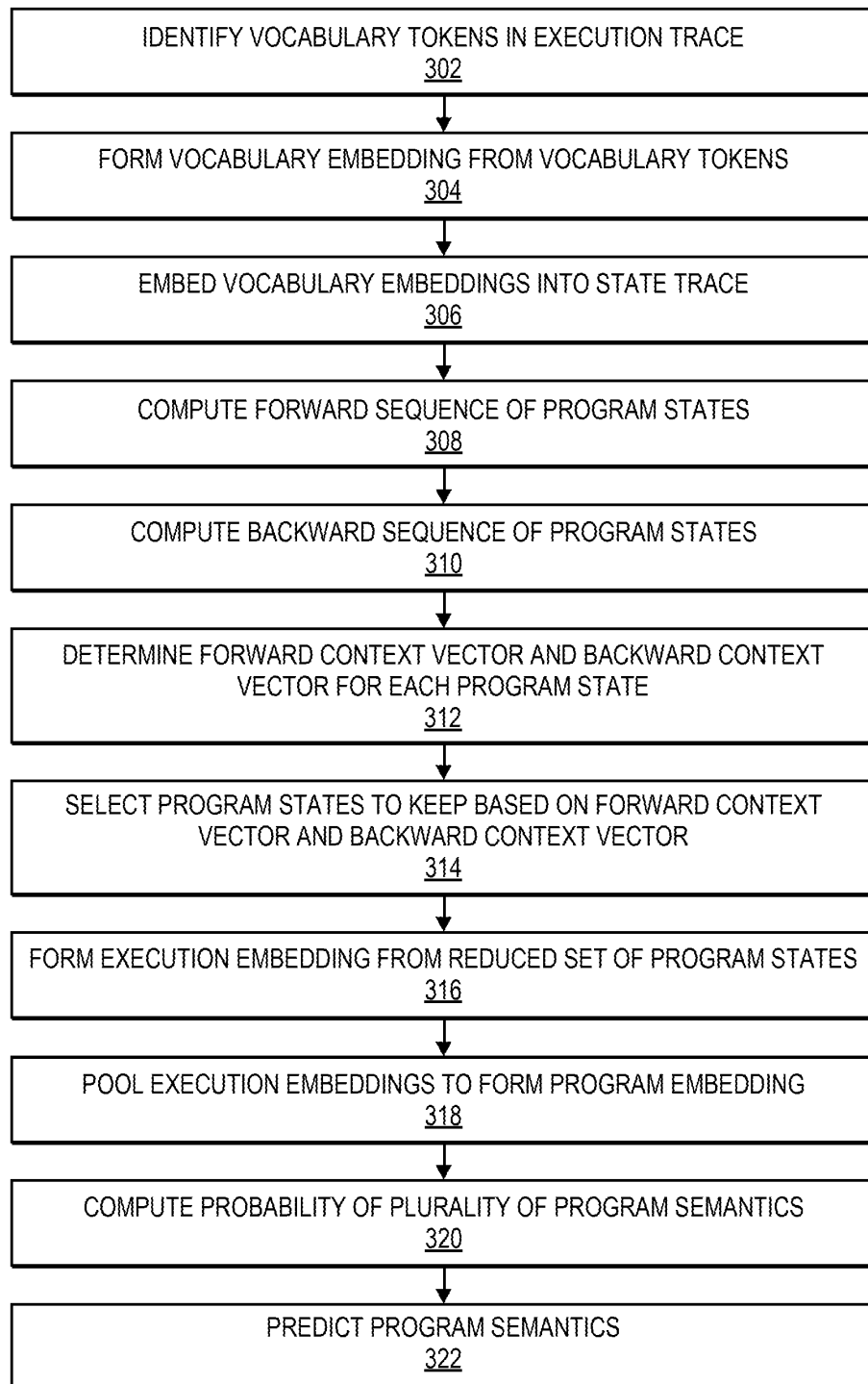


FIG. 1

2/10

**FIG. 2**

3/10

**FIG. 3**

4/10

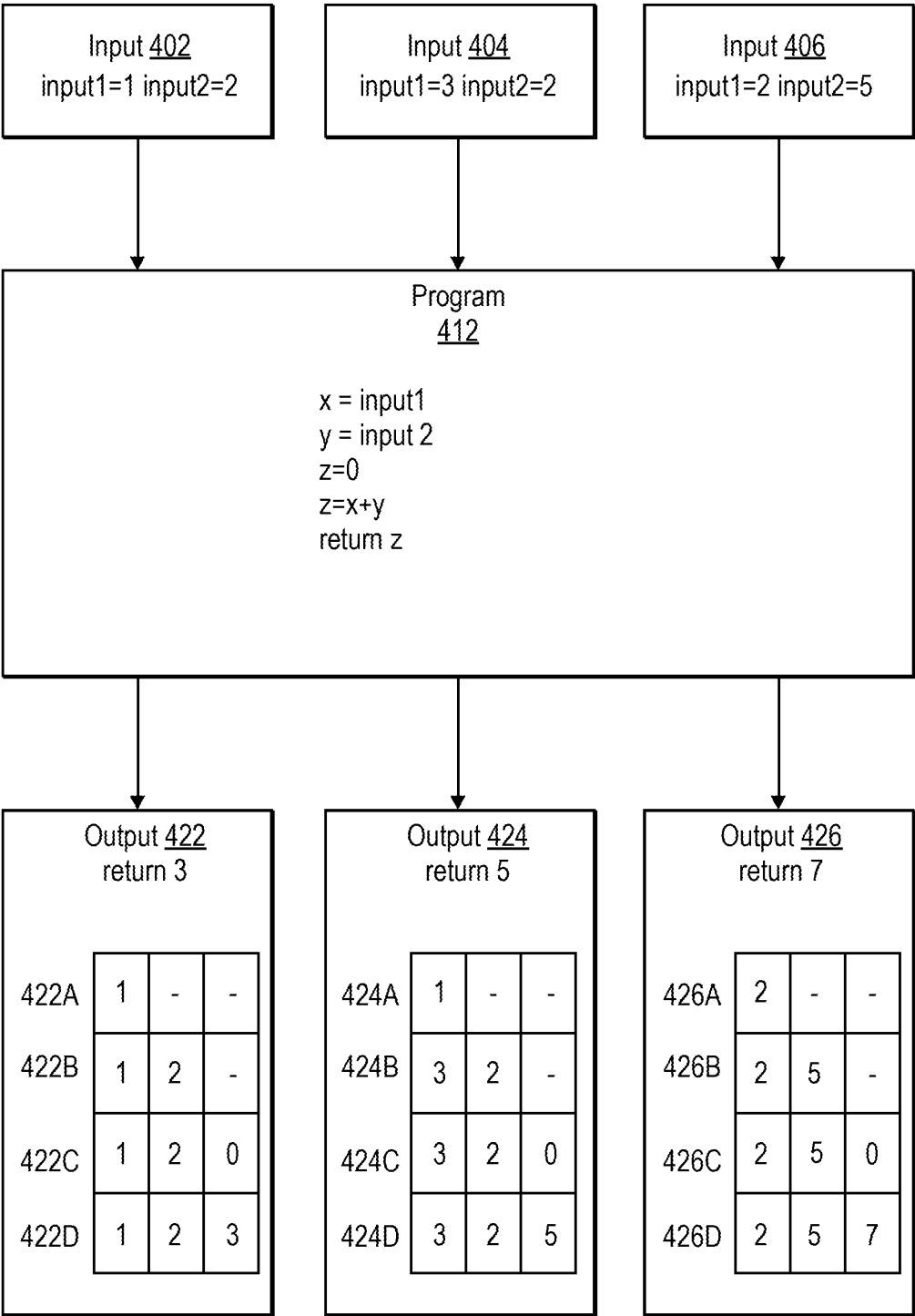


FIG. 4

5/10

| Tasks                         | Training Set | Validation Set | Training Set |
|-------------------------------|--------------|----------------|--------------|
| Array Average                 | 1,548        | 166            | 170          |
| Array k <sup>th</sup> largest | 561          | 168            | 162          |
| Array Max Diff                | 1,007        | 166            | 170          |
| Parenthesis                   | 723          | 168            | 165          |
| String Reverse                | 882          | 166            | 167          |
| Print Chessboard              | 5,279        | 166            | 166          |
| <b>Total</b>                  | <b>10K</b>   | <b>1K</b>      | <b>1K</b>    |

**FIG. 5A**

| Tasks                             | Training Set | Validation Set | Testing Set |
|-----------------------------------|--------------|----------------|-------------|
| Constant and Variable Propagation | 45,223       | 10,000         | 10,000      |
| Dead code Elimination             | 41,710       | 10,000         | 10,000      |
| Hoisting                          | 29,895       | 5,000          | 5,000       |
| Induction Variable Elimination    | 10,038       | 8,00           | 8,00        |
| Symbolic Execution Simulation     | 128,924      | 10,000         | 10,000      |

**FIG. 5B**

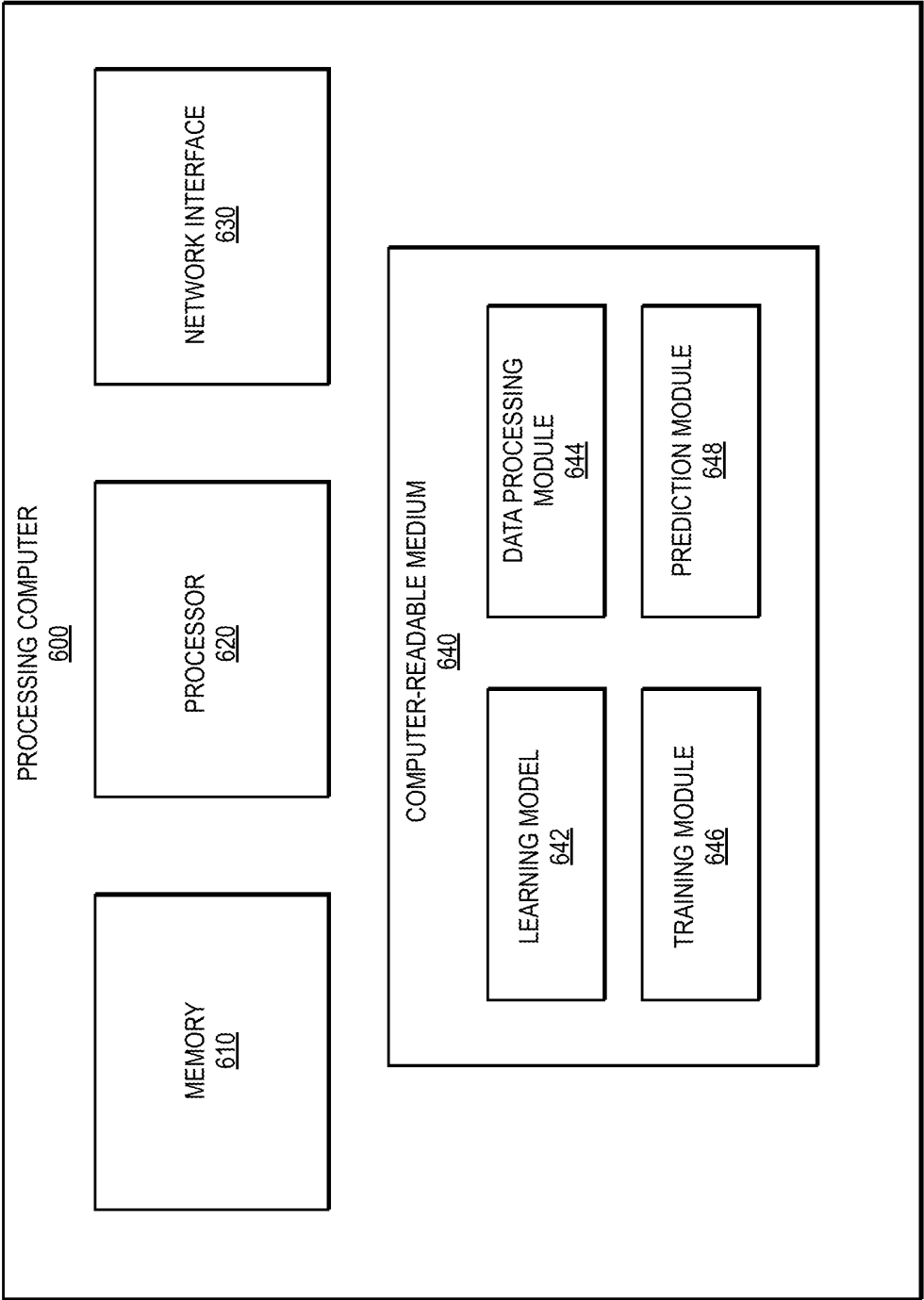
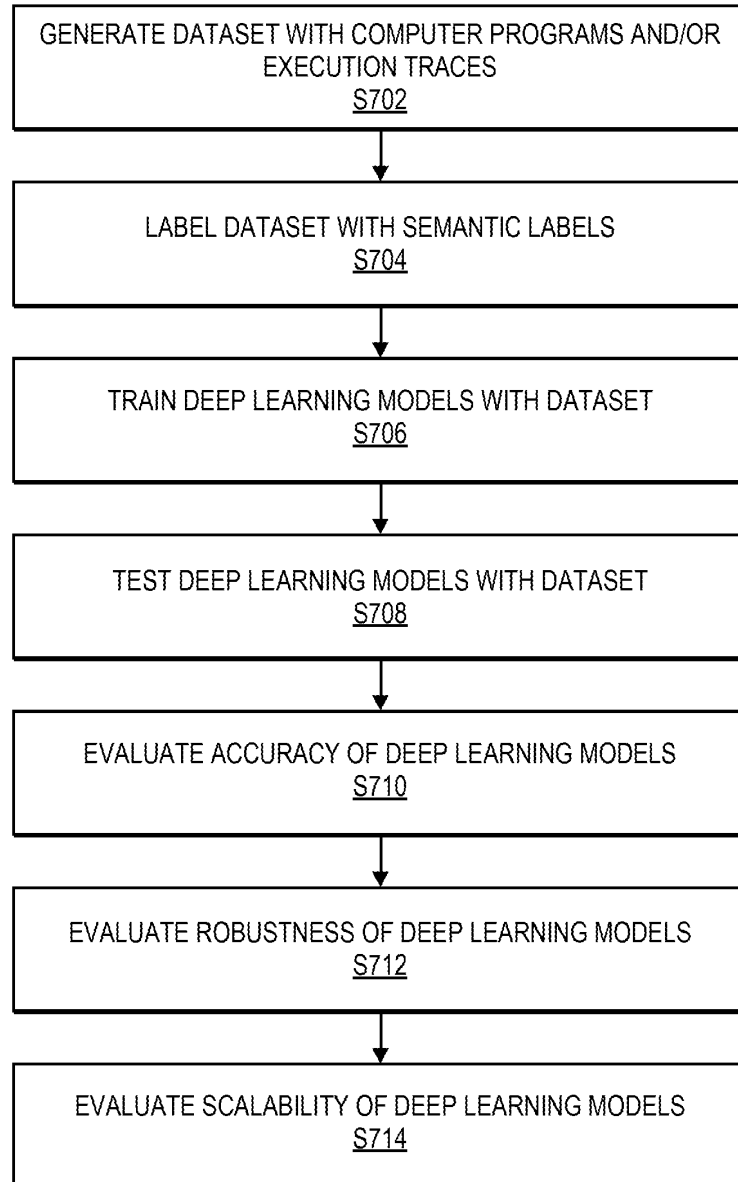


FIG. 6

7/10

**FIG. 7**

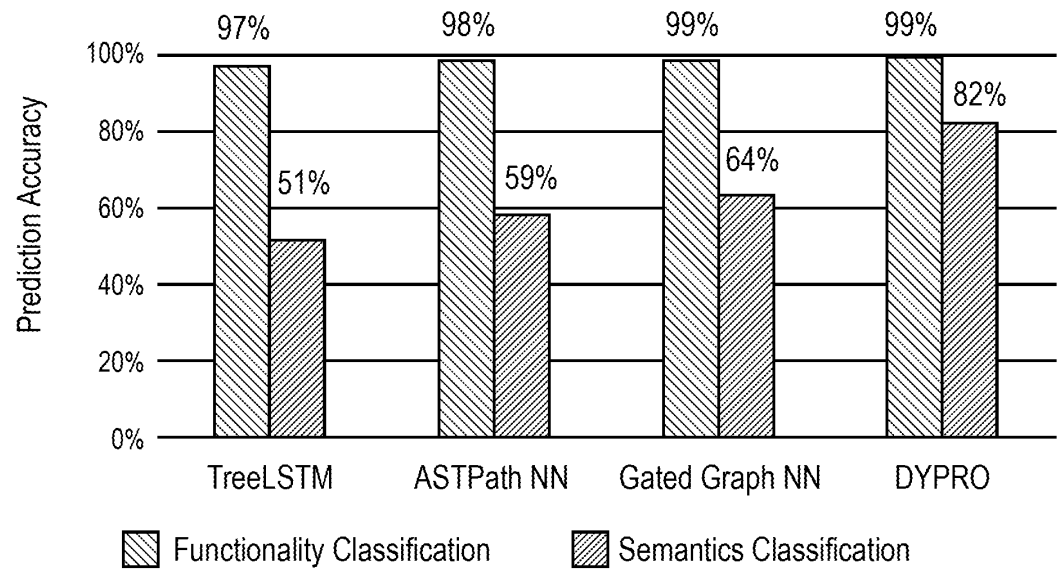


FIG. 8

9/10

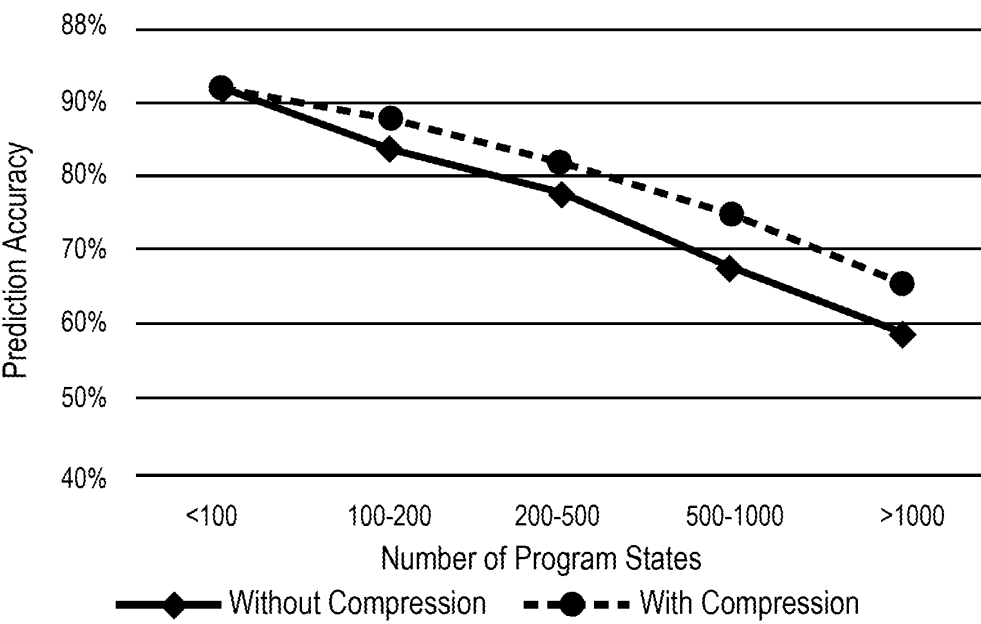


FIG. 9

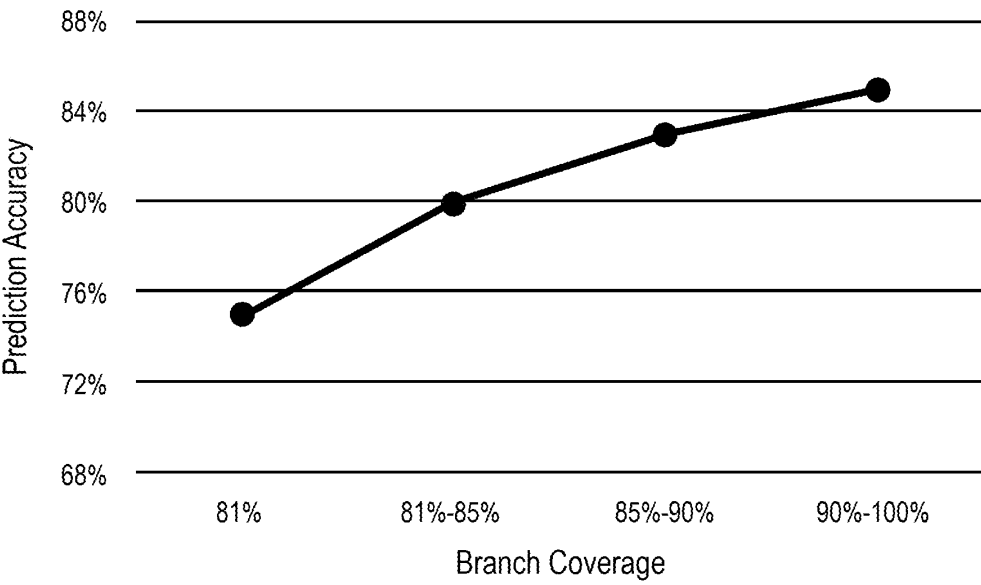


FIG. 10

| Models   | Constant and Variable Propagation |      | Dead code Elimination |      | Hoisting |      | Induction Variable Elimination |      |
|----------|-----------------------------------|------|-----------------------|------|----------|------|--------------------------------|------|
|          | Accuracy                          | F1   | Accuracy              | F1   | Accuracy | F1   | Accuracy                       | F1   |
| GGNN     | 98.1%                             | 0.92 | 91.5%                 | 0.87 | 63.8%    | 0.61 | 59.6%                          | 0.55 |
| TreeLSTM | 92.4%                             | 0.88 | 83.2%                 | 0.82 | 56.4%    | 0.59 | 53.5%                          | 0.52 |
| DyPro    | 98.3%                             | 0.95 | 94.8%                 | 0.91 | 92.7%    | 0.88 | 90.3%                          | 0.82 |

Result of semantic analysis.

FIG. 11