quant\_bits = 4, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 55.34% | 54.43% | 53.64% | 53.45% | 52.63% | 48.90% | 42.35% | 27.01% |
| Vgg16 | 64.84% | 65.20% | 64.47% | 63.48% | 63.06% | 61.76% | 58.46% | 48.34% |
| Resnet18 | 73.91% | 74.35% | 74.17% | 74.48% | 73.38% | 72.94% | 72.10% | 68.44% |
| Resnet34 | 74.72% | 74.74% | 74.80% | 74.38% | 73.46% | 73.34% | 72.80% | 69.85% |

quant\_bits = 4, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 83.45% | 82.80% | 82.42% | 81.84% | 81.56% | 78.94% | 73.83% | 57.63% |
| Vgg16 | 86.37% | 86.41% | 86.01% | 85.28% | 85.81% | 84.78% | 83.53% | 77.15% |
| Resnet18 | 92.49% | 92.28% | 92.72% | 92.43% | 92.10% | 92.44% | 92.00% | 91.21% |
| Resnet34 | 92.85% | 92.36% | 92.79% | 92.56% | 92.11% | 92.33% | 92.54% | 91.23% |

quant\_bits = 5, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 56.83% | 56.27% | 54.53% | 54.44% | 51.92% | 49.59% | 43.22% | 25.78% |
| Vgg16 | 65.85% | 64.88% | 64.93% | 64.59% | 64.06% | 62.88% | 58.95% | 48.85% |
| Resnet18 | 75.22% | 74.93% | 74.70% | 74.65% | 74.20% | 73.64% | 72.44% | 69.99% |
| Resnet34 | 75.35% | 74.67% | 75.03% | 74.04% | 73.83% | 73.71% | 73.34% | 71.15% |

quant\_bits = 5, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 83.92% | 83.66% | 83.11% | 82.97% | 81.51% | 79.73% | 74.35% | 56.12% |
| Vgg16 | 86.18% | 86.55% | 86.81% | 86.49% | 85.98% | 86.08% | 83.50% | 77.17% |
| Resnet18 | 92.98% | 92.91% | 93.01% | 92.55% | 92.58% | 92.65% | 92.48% | 92.11% |
| Resnet34 | 92.98% | 92.97% | 92.95% | 92.54% | 92.59% | 92.91% | 92.57% | 92.30% |

quant\_bits = 6, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 56.86% | 55.73% | 54.49% | 54.52% | 52.42% | 49.43% | 43.29% | 31.24% |
| Vgg16 | 65.50% | 65.18% | 64.97% | 64.43% | 63.03% | 62.74% | 60.83% | 47.81% |
| Resnet18 | 75.41% | 75.04% | 74.80% | 74.42% | 74.56% | 73.78% | 72.76% | 71.25% |
| Resnet34 | 71.50% | 74.91% | 74.31% | 74.86% | 74.81% | 74.08% | 73.12% | 71.78% |

quant\_bits = 6, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 84.08% | 83.89% | 83.43% | 82.40% | 81.66% | 79.64% | 74.70% | 62.89% |
| Vgg16 | 85.83% | 85.98% | 86.54% | 85.80% | 85.61% | 85.53% | 84.34% | 76.35% |
| Resnet18 | 93.10% | 93.40% | 93.09% | 93.03% | 93.16% | 92.71% | 92.70% | 92.09% |
| Resnet34 | 91.35% | 92.85% | 91.18% | 92.59% | 91.38% | 92.85% | 92.46% | 92.46% |

quant\_bits = 7, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 56.73% | 56.35% | 55.39% | 53.87% | 52.76% | 49.33% | 43.02% | 30.30% |
| Vgg16 | 65.30% | 65.73% | 64.86% | 64.60% | 63.40% | 62.51% | 58.37% | 48.75% |
| Resnet18 | 75.37% | 75.17% | 74.76% | 74.86% | 74.58% | 73.97% | 72.89% | 69.89% |
| Resnet34 | 75.02% | 74.10% | 72.76% | 74.79% | 73.40% | 74.24% | 68.59% | 71.56% |

quant\_bits = 7, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 84.06% | 83.81% | 83.91% | 82.87% | 81.99% | 79.76% | 73.96% | 61.97% |
| Vgg16 | 86.15% | 86.43% | 86.42% | 85.98% | 85.00% | 85.08% | 83.08% | 77.96% |
| Resnet18 | 93.03% | 92.88% | 92.92% | 92.91% | 93.06% | 92.94% | 92.71% | 91.51% |
| Resnet34 | 92.98% | 92.78% | 91.17% | 92.93% | 92.48% | 92.79% | 90.72% | 92.11% |

quant\_bits = 8, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 56.33% | 55.62% | 54.90% | 54.27% | 52.80% | 50.32% | 44.24% | 26.71% |
| Vgg16 | 65.66% | 65.21% | 64.86% | 64.12% | 63.65% | 62.32% | 59.89% | 47.95% |
| Resnet18 | 75.24% | 74.59% | 74.94% | 74.66% | 74.85% | 74.50% | 72.89% | 69.97% |
| Resnet34 | 75.66% | 75.11% | 69.14% | 75.13% | 74.65% | 72.91% | 73.43% | 71.75% |

quant\_bits = 8, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 84.11% | 83.53% | 83.07% | 82.54% | 81.95% | 80.12% | 75.34% | 56.72% |
| Vgg16 | 86.76% | 85.97% | 86.04% | 85.39% | 85.68% | 85.02% | 84.14% | 77.02% |
| Resnet18 | 93.17% | 93.04% | 92.84% | 92.81% | 92.89% | 93.19% | 92.80% | 91.84% |
| Resnet34 | 93.16% | 92.71% | 89.62% | 93.00% | 93.07% | 92.25% | 93.02% | 92.13% |

quant\_bits = False, top-1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 56.47% | 56.39% | 55.41% | 54.42% | 52.50% | 50.49% | 43.57% | 27.52% |
| Vgg16 | 65.10% | 65.41% | 64.39% | 64.32% | 63.52% | 63.25% | 59.51% | 52.00% |
| Resnet18 | 75.71% | 75.58% | 74.97% | 74.92% | 74.01% | 74.35% | 73.10% | 70.39% |
| Resnet34 | 75.82% | 75.10% | 75.65% | 75.46% | 74.94% | 73.98% | 74.50% | 72.51% |

quant\_bits = False, top-5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| modelname | 0.6 | 0.65 | 0.7 | 0.75 | 0.8 | 0.85 | 0.9 | 0.95 |
| Alexnet | 83.77% | 83.64% | 83.94% | 83.06% | 81.82% | 80.03% | 75.20% | 58.41% |
| Vgg16 | 86.70% | 87.03% | 86.50% | 86.36% | 86.04% | 86.24% | 84.69% | 80.94% |
| Resnet18 | 93.24% | 93.09% | 92.73% | 93.05% | 92.79% | 93.09% | 92.89% | 91.73% |
| Resnet34 | 93.34% | 92.91% | 93.20% | 93.04% | 92.69% | 92.96% | 92.89% | 92.63% |