

Figure 1: Performance vs Model Size
Task Specialization Outperforms Parameter Scaling

