

TABLE II
DOWNSTREAM TASKS AND DATASETS OF RAG

Task	Sub Task	Dataset	Method
QA	Single-hop	Natural Qustion(NQ) [111]	[26], [30], [34], [42], [45], [50], [52], [59], [64], [82] [3], [4], [22], [27], [40], [43], [54], [62], [71], [112] [20], [44], [72]
		TriviaQA(TQA) [113]	[13], [30], [34], [45], [50], [64] [4], [27], [59], [62], [112]
		SQuAD [114]	[22], [25], [43], [44], [71], [72] [20], [23], [30], [32], [45], [69], [112]
		Web Questions(WebQ) [115]	[3], [4], [13], [30], [50], [68]
		PopQA [116]	[7], [25], [67]
		MS MARCO [117]	[4], [40], [52]
	Multi-hop	HotpotQA [118]	[23], [26], [31], [34], [47], [51], [61], [82] [7], [14], [22], [27], [59], [62], [69], [71], [91]
		2WikiMultiHopQA [119]	[14], [24], [48], [59], [61], [91]
		MuSiQue [120]	[14], [51], [61], [91]
	Long-form QA	ELI5 [121]	[27], [34], [43], [49], [51]
		NarrativeQA(NQA) [122]	[45], [60], [63], [123]
		ASQA [124]	[24], [57]
		QMSum(QM) [125]	[60], [123]
	Domain QA	Qasper [126]	[60], [63]
		COVID-QA [127]	[35], [46]
		CMB [128],MMCU_Medical [129]	[81]
	Multi-Choice QA	QuALITY [130]	[60], [63]
		ARC [131]	[25], [67]
		CommonsenseQA [132]	[58], [66]
	Graph QA	GraphQA [84]	[84]
Dialog	Dialog Generation	Wizard of Wikipedia (WoW) [133]	[13], [27], [34], [42]
		KBP [134]	[74], [135]
		DuleMon [136]	[74]
	Task-oriented Dialog	CamRest [137]	[78], [79]
		Amazon(Toys,Sport,Beauty) [138]	[39], [40]
IE	Event Argument Extraction	WikiEvent [139]	[13], [27], [37], [42]
	Relation Extraction	RAMS [140]	[36], [37]
		T-REx [141],ZsRE [142]	[27], [51]
Reasoning	Commonsense Reasoning	HellaSwag [143]	[20], [66]
	CoT Reasoning	CoT Reasoning [144]	[27]
	Complex Reasoning	CSQA [145]	[55]
Others	Language Understanding	MMLU [146]	[7], [27], [28], [42], [43], [47], [72]
		WikiText-I03 [147]	[5], [29], [64], [71]
		StrategyQA [148]	[14], [24], [48], [51], [55], [58]
	Fact Checking/Verification	FEVER [149]	[4], [13], [27], [34], [42], [50]
		PubHealth [150]	[25], [67]
	Text Generation	Biography [151]	[67]
	Text Summarization	WikiASP [152]	[24]
		XSum [153]	[17]
	Text Classification	VioLens [154]	[19]
		TREC [155]	[33]
	Sentiment	SST-2 [156]	[20], [33], [38]
	Code Search	CodeSearchNet [157]	[76]
	Robustness Evaluation	NoMIRACL [56]	[56]
	Math	GSM8K [158]	[73]
	Machine Translation	JRC-Acquis [159]	[17]