

Microsoft-yn alymlary BitNet b1.58 2B4T - 2 milliard parametrli ilkinji 1-bitli emeli aň modelini işläp düzdüler, ol Apple M2 ýaly adaty prosessorlarda işläp bilýär. Tom's Hardware-iň habaryna görä, MIT-iň açyk ygtyýarlygy bilen çykan model Hugging Face platformasynda elýeterli. Onuň aýratynlygy 1-bit agyrlyk derejesini ulanyp (-1, 0, +1), bu adatdaky 16 ýa-da 32-bit formatlaryny ulanýan modeller bilen deňeşdirilende, ýatda saklamaga we hasaplaýyş çeşmelere bolan talaplaryny ep-esli azaldýar.

BitNet, 4 trillion tekst birliklerinde (tokenda) öwredilip, Google Gemma 3 1B ýaly modeller bilen öndürijilik boýunça deňeşdirilýär, ýöne bary-ýogy 400 MB ýat sarp edýär - bäsdeşlerden birnäçe esse az. Bu, kuwwaty çäklendirilen enjamlar üçin ony amatly edýär we kuwwatly grafiki kartalara (GPU) ýa-da nerw proseslerine (NPU) garaşlylygy peseldýär. Emma iň ýokary öndürijilik üçin, GPU-ni häzirlilikçe goldamaýan ýörite bitnet.cpp çarçuwasy gerek.