



Jensen Huang a role NVIDIA v revoluci hlubokého učení



"AI won't take your job: The person who uses ai will.." -
Jensen Huang

Od her k umělé inteligenci: Jak NVIDIA a Jensen Huang změnili svět

Málokterá firma ztělesňuje současnou revoluci v umělé inteligenci tak jako NVIDIA [3]. A málokterý technologický lídr měl tak zásadní vliv na její směřování jako Jensen Huang, její spoluzakladatel a dlouholetý CEO. Příběh NVIDIE je fascinující ukázkou toho, jak se z výrobce specializovaného hardwaru pro hráče videoher stala páteří infrastrukturu pro jednu z největších technologických transformací v historii [4]. Vše začalo v roce 1993, kdy byla hlavním cílem NVIDIE produkce stále výkonnějších grafických procesorů (GPU), které by hráčům nabídly co nejrealističtější 3D grafiku. Klíčový zlom však přišel s pochopením, že masivně paralelní architektura těchto čipů je jako stvořená i pro jiné úkoly. Zatímco klasické procesory (CPU) jsou skvělé v řešení několika složitých úkolů za sebou, GPU dokážou zpracovávat tisíce jednodušších operací současně. A právě to je přesně to, co vyžadují neuronové sítě. To, co bylo původně navrženo pro plynulé vykreslování textur a světél ve hrách, se ukázalo jako ideální nástroj pro náročné maticové výpočty, které jsou jádrem hlubokého učení.

Definitivním potvrzením tohoto směřování se stal rok 2012 (30. září). Tehdy neuronová síť AlexNet, trénovaná právě na grafických kartách NVIDIA GeForce, s drtivou převahou zvítězila v prestižní soutěži ImageNet. Pro svět umělé inteligence to byl zlomový okamžik. Ukázalo se, že díky GPU je možné trénovat modely s dříve nemyslitelnou složitostí a rychlostí. Proces, který by na běžných procesorech trval týdny, byl najednou hotový za několik dní.

Samotný úspěch neleží jen v geniálním hardwaru ale i geniálního softwaru. Tím je softwarová platforma CUDA (Compute Unified Device Architecture). CUDA představovala a představuje revoluční krok, protože programátorům nabízí relativně jednoduchý způsob, jak využít výpočetní sílu GPU pro jakýkoliv účel, nejen pro grafiku. NVIDIA tak vytvořila kolem svého hardwaru nepostradatelný ekosystém, který jí dneska přinesl slávu a umožnil jí se stát jednou z největších společností na světě.

Pro programátory, kteří dnes v drtivé většině pracují v Pythonu, se tento ekosystém stal standardem. NVIDIA neusnula na vavřínech a vyvinula specializované knihovny, které celý proces ještě zjednodušují a zrychlují. Klíčovou je knihovna cuDNN (CUDA Deep Neural Network), která poskytuje vysoce optimalizované implementace základních stavebních bloků neuronových sítí,

jako jsou konvoluce nebo aktivační funkce [2]. Prakticky všechny moderní frameworky pro hluboké učení, jako TensorFlow nebo PyTorch, se na cuDNN spoléhají, aby dosáhly maximálního výkonu.

Když je model natrénovaný, přichází na řadu jeho nasazení do praxe – takzvaná inference. I na to má NVIDIA odpověď v podobě TensorRT, nástroje pro optimalizaci neuronových sítí, který zajišťuje co nejrychlejší odezvu a nejvyšší propustnost v reálných aplikacích, od autonomních vozidel po rozpoznávání řeči. NVIDIA jde dokonce ještě dál s projekty jako RAPIDS, což je sada knihoven, která umožňuje provádět celý proces datové analýzy – od načtení dat přes jejich zpracování až po vizualizaci – kompletně na GPU. Nástroje jako cuDF napodobují populární knihovnu Pandas, což datovým vědcům umožňuje akcelarovat jejich stávající kód s minimálními úpravami.

Jensen Huang

Za fenomenálním úspěchem NVIDIE nestojí jen technologická vize, ale především osobnost a odhodlání jejího zakladatele. Jensen Huang, který se narodil na Tchaj-wanu a do Spojených států se s rodinou přestěhoval v dětství, je známý svou ikonickou černou koženou bundou, ale především svou ochotou podstupovat obrovská rizika. Počátky firmy nebyly jednoduché. V roce 1996, pouhé tři roky po založení, stála NVIDIA jen 30 dní od platební neschopnosti. Její první čip NV1 nebyl komerčně úspěšný a firma musela propustit více než polovinu zaměstnanců. Místo toho, aby se Jensen Huang držel původního plánu, udělal radikální rozhodnutí: vsadil vše na zcela novou architekturu, která se soustředila výhradně na trojúhelníkovou rasterizaci – techniku, která se později stala standardem v 3D grafice. Tento riskantní krok se vyplatil s příchodem čipu RIVA 128, který NVIDII zachránil a katapultoval ji mezi špičku v oboru. Tato zkušenost hluboce ovlivnila firemní kulturu. Ve firmě se zavedl třeba princip "intelektuální poctivosti" (Intellectual Honesty), který nabádá zaměstnance, aby rychle přiznali, pokud se nějaký projekt nevyvíjí správně, a bez obav změnil směr. Dále se třeba zakladatel a firma nekonvenčně staví k individuálním schůzkám, kterým se vyhýbá a podporuje plochou organizační strukturu, aby se podpořila přímá komunikace a kolektivní řešení problémů.

Co se Huangovi vize týče, zakladatel NVIDIE mluví o budoucnosti, kde budou digitální a biologičtí pracovníci spolupracovat. Předpovídá, že umělá inteligence dosáhne obecné inteligence (AGI), která se vyrovná nebo překoná

lidské schopnosti, a to možná už do pěti let [1]. Společnost jako taková masivně investuje do autonomních systémů, robotiky, digitálních dvojčat a platformě Omniverse - v podstatě je to operační systém pro metaverzum a tvorbu digitálních dvojčat¹.

Příběh Jensena Huanga a NVIDIE je tak mnohem víc než jen příběhem o rychlých čipech. NVIDIA neprodává jen grafické karty; prodává přístup k superpočítačovému výkonu a poskytuje nástroje, které tuto sílu zpřístupňují celému světu. Tím nejenže definovala současnou éru umělé inteligence, ale položila i základy pro to, co teprve přijde [3].

¹Jde o virtuální, fyzikálně přesnou repliku reálného objektu, procesu nebo celého prostředí (např. továrny, města, motoru).

Použité zdroje

Reference

- [1] Jordan Novet. Nvidia ceo says ai could be 'fairly competitive' with humans in 5 years. CNBC, November 2023. URL: <https://www.cnbc.com/2023/11/29/nvidia-ceo-ai-will-be-fairly-competitive-with-humans-in-5-years.html>.
- [2] NVIDIA Corporation. Nvidia cudnn, 2024. URL: <https://developer.nvidia.com/cudnn>.
- [3] Anton Shilov. Nvidia gtc 2024: Why nvidia dominates ai. *EE Times*, March 2024. URL: <https://www.eetimes.com/nvidia-gtc-2024-why-nvidia-dominates-ai/>.
- [4] Wikipedia contributors. Nvidia — Wikipedia, The Free Encyclopedia, 2024. URL: <https://en.wikipedia.org/wiki/Nvidia>.