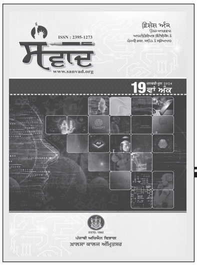

ਸੰਵਾਦ-19 (ਜਨਵਰੀ-ਜੂਨ 2024) ISSN : 2395-1273 www.sanvad.org

ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਦੇ ਸਮਾਜਕ ਸੰਤੁਲਨ ਨੂੰ ਵਿਗਾੜਨ ਵਾਲੇ ਖ਼ਤਰੇ

ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਕੀ ਹੈ?

AI ਦਾ ਮਤਲਬ ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਹੈ, ਜੋ ਕੰਪਿਊਟਰ ਜਾਂ ਮਸ਼ੀਨ ਦੀ ਉਹਨਾਂ ਕੰਮਾਂ ਨੂੰ ਕਰਨ ਦੀ ਯੋਗਤਾ ਹੈ ਜਿਨ੍ਹਾਂ ਲਈ ਆਮ ਤੌਰ ’ਤੇ ਮਨੁੱਖੀ ਬੁੱਧੀ ਦੀ ਲੋੜ ਹੁੰਦੀ ਹੈ, ਜਿਵੇਂ ਕਿ ਭਾਸ਼ਾ ਨੂੰ ਸਮਝਣਾ, ਚਿੱਤਰਾਂ ਨੂੰ ਪਛਾਣਨਾ, ਫੈਸਲੇ ਲੈਣਾ ਅਤੇ ਤਜ਼ਰਬੇ ਤੋਂ ਸਿੱਖਣਾ। AI ਇਕ ਵਿਆਪਕ ਸ਼ਬਦ ਹੈ ਜੋ ਬਹੁਤ ਸਾਰੀਆਂ ਵੱਖ-ਵੱਖ ਤਕਨਾਲੋਜੀਆਂ ਨੂੰ ਸ਼ਾਮਲ ਕਰਦਾ ਹੈ, ਜਿਵੇਂ ਕਿ ਮਸ਼ੀਨ-ਲਰਨਿੰਗ (ਸਿਖਲਾਈ), ਡੂੰਘੀ ਸਿੱਖਿਆ (deep learning), ਕੁਦਰਤੀ ਭਾਸ਼ਾ ਪ੍ਰੋਸੈਸਿੰਗ, ਕੰਪਿਊਟਰ ਰਾਹੀਂ ਦ੍ਰਿਸ਼ਟੀ ਪਛਾਣ, ਭਾਸ਼ਣ (ਬੋਲ) ਪਛਾਣ, ਅਤੇ ਹੋਰ। AI ਦੀ ਵਰਤੋਂ ਵੱਖ-ਵੱਖ ਉਦੇਸ਼ਾਂ ਲਈ ਕੀਤੀ ਜਾ ਸਕਦੀ ਹੈ, ਜਿਵੇਂ ਕਿ ਮਨੋਰੰਜਨ, ਸਿੱਖਿਆ, ਸਿਹਤ ਸੰਭਾਲ, ਕਾਰੋਬਾਰ, ਸੁਰੱਖਿਆ, ਅਤੇ ਹੋਰ। AI ਕੰਪਿਊਟਰ ਵਿਗਿਆਨ ਵਿਚ ਅਧਿਐਨ ਦਾ ਇਕ ਖੇਤਰ ਇਹ ਵੀ ਹੈ ਜਿਸਦਾ ਉਦੇਸ਼ ਬੁੱਧੀਮਾਨ ਮਸ਼ੀਨਾਂ ਨੂੰ ਬਣਾਉਣਾ ਅਤੇ ਸਮਝਣਾ ਹੈ। ਦਰਅਸਲ AI ਦਾ ਵਿਸ਼ਾ ਬਹੁਤ ਵਿਸ਼ਾਲ ਹੈ ਤੇ ਅਕਸਰ ਲੋਕ ਇਸ ਦੇ ਵੱਖ-ਵੱਖ ਮਤਲਬ ਕੱਢ ਲੈਂਦੇ ਹਨ। ਆਮ ਤੌਰ ’ਤੇ ਜੋ ਕੋਈ ਵੀ ਦਿਮਾਗੀ ਤੇ ਗੁੰਝਲਦਾਰ ਕੰਮ ਇਨਸਾਨ ਕਰਦਾ ਹੈ ਜੇ ਉਹ ਮਸ਼ੀਨ ਰਾਹੀਂ ਇਨਸਾਨ ਨਾਲੋਂ ਚੰਗਾ ਕੀਤਾ ਜਾਵੇ ਤਾਂ ਉਸ ਨੂੰ AI ਕਿਹਾ ਜਾਂਦਾ ਹੈ। ਅਸਲੀਅਤ ਵਿਚ ਆਧੁਨਿਕ AI ਬਹੁਤ ਅੱਗੇ ਨਿਕਲ ਚੁੱਕੀ ਹੈ ਤੇ ਬਹੁਤ ਤੇਜ਼ੀ ਨਾਲ ਤਰੱਕੀ ਕਰ ਰਹੀ ਹੈ। ਇਥੋਂ ਤਕ ਕਿ AI ਵਿਗਿਆਨੀ ਵੀ ਨਹੀਂ ਜਾਣਦੇ ਕਿ ਇਸ ਦੀ ਸ਼ਕਤੀ ਕਿਤਨੀ ਪ੍ਰਬਲ ਹੋਵੇਗੀ।

ਡਾ. ਕੁਲਬੀਰ ਸਿੰਘ ਥਿੰਦ
ਪਿਤਾਮਾ
ਪੰਜਾਬੀ ਫੌਂਟਸ
ਕੈਲੀਫੋਰਨੀਆ
ਅਮਰੀਕਾ
+1 (510) 386-5033

ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਕਿਵੇਂ ਕੰਮ ਕਰਦੀ ਹੈ?

ਇਸ ਵਿਸ਼ੇ ਨੂੰ ਵਿਸਤਾਰ ਨਾਲ ਸਮਝਾਉਣਾ ਇਸ ਲੇਖ ਦਾ ਮਕਸਦ ਨਹੀਂ ਹੈ। ਇਸ ਬਾਰੇ ਜਾਣਕਾਰੀ ਵਿਕੀਪੀਡਿਆ (wikipedia.org) ਤੋਂ ਮਿਲ ਸਕਦੀ ਹੈ। ਵੈਸੇ ਵੀ ਆਮ ਮਨੁੱਖਾਂ ਲਈ ਇਸ ਨੂੰ ਸਮਝਣਾ ਔਖਾ ਹੈ। AI ਤਾਂ ਸੰਭਵ ਹੀ ਇਸ ਲਈ ਹੋਈ ਹੈ ਕਿਉਂ ਕਿ ਕੰਪਿਊਟਰ ਬਹੁਤ ਤਾਕਤਵਰ ਹੋ ਗਏ ਹਨ ਅਤੇ ਬਹੁਤ ਵੱਡੇ ਡਾਟਾ-ਸੈਟ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) ਉਪਲੱਬਧ ਹੋ ਗਏ ਹਨ। ਇਸ ਤੋਂ ਇਲਾਵਾ AI ਵਿਗਿਆਨਕਾਂ ਨੂੰ AI ਦੇ ਕੰਪਿਊਟਰਾਂ ਨੂੰ ਸਹੀ ਤਰ੍ਹਾਂ ਪ੍ਰੋਗਰਾਮ ਕਰਨਾ ਆ ਗਿਆ ਹੈ। ਅਤੇ AI ਨੂੰ ਬੇਹਤਰ ਕਰਨ ਲਈ ਇਸ ਪਿੱਛੇ ਕਈ ਹੋਰ ਇਨਸਾਨਾਂ ਦਾ ਕੰਮ ਹੁੰਦਾ ਹੈ ਜਿਨ੍ਹਾਂ ਦਾ ਵਿਸ਼ਾ ਹੁੰਦਾ ਹੈ AI ਵੱਲੋਂ ਕਿਸੇ ਸਵਾਲ ਦੇ ਉਤਰ ਨੂੰ ਸਹੀ ਕਰਨ ਲਈ ਸੁਝਾਅ ਦੇਣਾ।

ਹੁਣ ਤਾਈਂ ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਦੀ ਵਰਤੋਂ ਤਰਕ (reasoning) ਵਾਸਤੇ, ਸਮੱਸਿਆ ਹੱਲ ਕਰਨ ਕਈ, ਗਿਆਨ ਦੀ ਪੇਸ਼ਗੀ ਲਈ, ਕੁਦਰਤੀ ਭਾਸ਼ਾ ਪ੍ਰੋਸੈਸਿੰਗ ਲਈ, ਅਨੁਭਵ ਬੇਹਤਰ ਕਰਨ ਲਈ,

ਰੋਬੋਟਿਕਸ ਦੀ ਵਰਤੋਂ ਵਾਸਤੇ, ਸਮਾਜਕ/ਸੋਸ਼ਲ ਇੰਟੈਲੀਜੈਂਸ ਸਮਝਣ ਤੇ ਵਰਤੋਂ ਵਾਸਤੇ, ਆਮ/ਜਨਰਲ ਇੰਟੈਲੀਜੈਂਸ ਅਤੇ ਚਿੱਤਰ ਵਿਸ਼ਲੇਸ਼ਣ ਆਦਿ ਲਈ ਕੀਤੀ ਜਾ ਰਹੀ ਹੈ। ਵਰਤਮਾਨ ਵਿਚ, AI ਕੁਦਰਤੀ ਭਾਸ਼ਾ ਪ੍ਰੋਸੇਸਿੰਗ ਅਤੇ ਇਸ ਨੂੰ ਸਮਝਣ ਵਿਚ ਬਹੁਤ ਵਧੀਆ ਹੋ ਗਈ ਹੈ, ਹੁਣ ਤਕ ਤਾਂ ਇਹ ਪ੍ਰਸ਼ਨਾਂ ਨੂੰ ਸਮਝਦੀ ਹੈ ਅਤੇ ਉਚਿਤ ਜਵਾਬ ਪ੍ਰਦਾਨ ਕਰਦੀ ਹੈ। ਅਤੇ ਇਹ ਹਰ ਦਿਨ ਹੋਰ ਬਿਹਤਰ ਹੋ ਰਹੀ ਹੈ। ਕਮਾਲ ਦੇ ਰੋਬੋਟ ਵੀ ਬਣ ਗਏ ਹਨ ਜੋ ਹਰ-ਰੋਜ਼ ਬੇਹਤਰ ਹੁੰਦੇ ਜਾ ਰਹੇ ਹਨ। AI ਕਈ ਹੋਰ ਕੰਮਾਂ ਲਈ ਵਰਤਨਯੋਗ ਹੋ ਗਈ ਹੈ ਜਿਵੇਂ ਕਿ ਛਾਤੀ ਐਕਸ-ਰੇ ਜਾਂ ਮੈਮੋਂਗਰਾਮ ਦਾ ਸਹੀ ਨਿਚੋੜ ਕਰਨਾ।

ਕੁਦਰਤੀ ਵਿਕਾਸ ਦੁਆਰਾ (through evolution), ਜ਼ਿਆਦਾਤਰ ਜੀਵ ਆਪਣੇ ਜੀਵਤ ਰਹਿਣ ਲਈ ਅਤੇ ਪ੍ਰਸਾਰ (ਸੰਤਾਨ ਪੈਦਾ) ਕਰਨ ਲਈ ਬਣੇ ਹਨ। ਕੁਝ ਜਾਨਵਰਾਂ ਨੇ ਕਬੀਲਿਆਂ ਵਿਚ ਰਹਿਣਾ ਸਿਖ ਲਿਆ ਤੇ ਸਮਾਜ ਨੂੰ ਦਰਜਾ-ਬੰਦ ਕਰਨਾ ਸਿਖ ਲਿਆ। ਪਰ ਮਨੁੱਖ ਦੂਜੇ ਜੀਵਾਂ ਨਾਲੋਂ ਬਹੁਤ ਵੱਖਰੇ ਹਨ। ਭਾਸ਼ਾ ਸਿੱਖਣ ਕਾਰਨ ਮਨੁੱਖ ਹੋਰ ਮਨੁੱਖਾਂ ਅਤੇ ਆਉਣ ਵਾਲੀਆਂ ਪੀੜ੍ਹੀਆਂ ਨੂੰ ਗਿਆਨ ਦੇਣ ਦੇ ਯੋਗ ਹਨ। ਆਧੁਨਿਕ ਮਨੁੱਖ ਛੋਟੇ ਕਬੀਲਿਆਂ ਤੋਂ ਨਿਕਲ ਕੇ ਵੱਡੇ ਸਮਾਜਕ ਸਮੂਹਾਂ ਵਿਚ ਵੰਡਿਆ ਹੋਇਆ ਹੈ ਜਿਸ ਵਿਚ ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਦਰਜਾਬੰਦੀ ਹੈ। ਮਨੁੱਖੀ ਸਮਾਜ ਦੇ ਕ੍ਰਿਆ-ਕ੍ਰਮ ਕੇਵਲ ਜੀਵਤ ਰਹਿਣ ਲਈ ਅਤੇ ਸੰਤਾਨ ਪੈਦਾ ਕਰਨ ਤਕ ਸੀਮਤ ਨਹੀਂ ਹਨ। ਇਨਸਾਨ ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕਲਾ ਤੇ ਗਿਆਨ ਦੇ ਮਾਹਰ ਹਨ ਜੋ ਕੇਵਲ ਜੀਵਤ ਰਹਿਣ ਲਈ ਜ਼ਰੂਰੀ ਨਹੀਂ। ਭਾਵੇਂ ਮਨੁੱਖਾਂ ਕੋਲ ਵਿਕਸਤ ਦਿਮਾਗ ਹਨ ਜੋ ਲਗਭਗ 100 ਬਿਲੀਅਨ ਨਿਊਰੋਨਜ਼ ਦੇ ਨਾਲ ਬਣੇ ਹਨ ਅਤੇ ਜਿਨ੍ਹਾਂ ਦੇ ਅਗੇ 100 ਟ੍ਰਿਲੀਅਨ ਕੁਨੈਕਸ਼ਨ (ਜੋੜ) ਹੁੰਦੇ ਹਨ, ਪਰ ਵਿਅਕਤੀਗਤ ਤੌਰ 'ਤੇ ਅਸੀਂ ਜੋ ਸਿੱਖਦੇ ਹਾਂ ਉਸ ਦੀਆਂ ਸੀਮਾਵਾਂ ਹਨ, ਅਤੇ ਅਸੀਂ ਗਿਆਨ ਨੂੰ ਹੋਰ ਮਨੁੱਖਾਂ ਨੂੰ ਤੇਜ਼ੀ ਨਾਲ ਨਹੀਂ ਦੇ ਸਕਦੇ। ਸਾਡੇ ਕੋਲ ਸੰਵੇਦਨਾਸ਼ੀਲ ਅੰਗ ਵੀ ਹਨ ਜੋ ਸਾਨੂੰ ਸਾਡੇ ਵਾਤਾਵਰਣ ਬਾਰੇ ਜਾਣਕਾਰੀ ਪ੍ਰਦਾਨ ਕਰਦੇ ਹਨ, ਪਰ ਇਨ੍ਹਾਂ ਦੀਆਂ ਵੀ ਆਪਣੀਆਂ ਸੀਮਾਵਾਂ ਹਨ। ਮਨੁੱਖ ਅੰਦਰੂਨੀ ਭਾਵਨਾਵਾਂ (emotions) ਦੇ ਅਧੀਨ ਵੀ ਹੁੰਦੇ ਹਨ। ਕਈ ਹਾਲਾਤਾਂ ਵਿਚ ਮਨੁੱਖੀ ਦਿਮਾਗ ਸਹੀ ਕੰਮ ਵੀ ਨਹੀਂ ਕਰਦੇ। ਦੂਜੇ ਪਾਸੇ ਮਸ਼ੀਨਾਂ ਵੱਡੀ ਮਾਤਰਾ ਵਿਚ ਗਿਆਨ ਇਕੱਠਾ ਕਰ ਸਕਦੀਆਂ ਹਨ ਅਤੇ ਉਸ ਗਿਆਨ ਨੂੰ ਪ੍ਰਾਪਤ ਕਰਨ ਤੋਂ ਬਾਅਦ ਅੱਗੇ ਵਿਸ਼ਲੇਸ਼ਣ ਕਰ ਕੇ ਹੋਰ ਸੂਝਵਾਨ ਬਣ ਸਕਦੀਆਂ ਹਨ ਅਤੇ ਉਨ੍ਹਾਂ ਦੇ ਸੂਝਵਾਨ ਹੋਣ ਦੀ ਸਮਰਥਾ ਵਕਤ ਤੇ ਵਰਤੋਂ ਨਾਲ ਵਧਦੀ ਹੀ ਰਹਿੰਦੀ ਹੈ। ਕੋਈ ਵੀ ਇਨਸਾਨ ਸਾਰੇ ਜਹਾਨ ਦੇ ਮਨੁੱਖਾਂ ਬਾਰੇ ਜਾਣਕਾਰੀ ਨਹੀਂ ਰੱਖ ਸਕਦਾ, ਪਰ AI ਮਸ਼ੀਨਾਂ ਇੰਜ ਕਰ ਸਕਦੀਆਂ ਹਨ ਤੇ ਇੰਟਰਨੈੱਟ ਨੂੰ ਫਰੋਲ ਕੇ ਉਹਨਾਂ ਬਾਰੇ ਬੇਹੱਦ ਜਾਣਕਾਰੀ ਇਕੱਠੀ ਕਰ ਸਕਦੀਆਂ ਹਨ। ਇਸ ਜਾਣਕਾਰੀ ਵਿਚ ਹਰ ਤਰ੍ਹਾਂ ਦੀ ਜਾਣਕਾਰੀ ਮੌਜੂਦ ਹੈ ਜੋ ਤੁਸੀਂ ਸੋਚ ਵੀ ਨਹੀਂ ਸਕਦੇ। ਐਸੀਆਂ ਮਸ਼ੀਨਾਂ ਜਾਣ ਸਕਦੀਆਂ ਹਨ ਕਿ ਕੋਈ ਇਨਸਾਨ ਕਿਥੋਂ ਦਾ ਹੈ, ਕੀ ਕੰਮ ਕਰਦਾ ਹੈ, ਉਸ ਦੇ ਕਿਹੋ ਜਿਹੇ ਖਿਆਲ ਹਨ, ਉਹ ਕਿਸ ਪਾਰਟੀ ਦਾ ਹਮਾਇਤੀ ਹੈ ਅਤੇ ਕੈਸੀਆਂ ਚੀਜ਼ਾਂ ਖਰੀਦਦਾ ਹੈ, ਉਸ ਦੇ ਫੋਨ ਨੰਬਰ ਕੀ ਹਨ, ਈ-ਮੇਲ ਕੀ ਹੈ, ਆਦਿ। ਇਥੋਂ ਤਕ ਕਿ ਐਸੀ ਮਸ਼ੀਨ ਪਾਸ-ਵਰਡ ਵੀ ਪਤਾ ਲਾ ਸਕਦੀ ਹੈ। ਇਸ ਤੋਂ ਇਲਾਵਾ, ਇਕ ਐਸੀ ਮਸ਼ੀਨ ਤੇਜ਼ੀ ਨਾਲ ਗਿਆਨ ਨੂੰ ਕਿਸੇ ਹੋਰ ਮਸ਼ੀਨ ਨੂੰ ਦੇ ਸਕਦੀ ਹੈ। ਇਸ ਤਰ੍ਹਾਂ, AI ਦੀਆਂ ਸੰਭਾਵਨਾਵਾਂ ਬੇਅੰਤ ਹਨ। ਸਭ ਤੋਂ ਵੱਡੀ ਗਲ ਇਹ ਹੈ ਕਿ ਇਸ ਦੇ ਹੋਰ ਵਿਕਾਸ ਦੀ ਦਰ ਬਹੁਤ ਹੈ।

ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਦੇ ਕੁਝ ਖਾਸ ਖ਼ਤਰੇ ਅਤੇ ਨੁਕਸਾਨ

1. ਜਾਣਕਾਰੀ ਦੀ ਪਰਦੇਦਾਰੀ ਅਤੇ ਨਕਲੀ ਬੁੱਧੀ ਅਤੇ ਕਾਪੀਰਾਈਟ

ਮਸ਼ੀਨ-ਲਰਨਿੰਗ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਨੂੰ ਵੱਡੀ ਮਾਤਰਾ ਵਿਚ ਡਾਟਾ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) ਦੀ ਲੋੜ ਹੁੰਦੀ ਹੈ। ਇਸ ਡਾਟਾ ਨੂੰ ਪ੍ਰਾਪਤ ਕਰਨ ਲਈ ਵਰਤੀਆਂ ਗਈਆਂ ਤਕਨੀਕਾਂ ਨੇ ਪਰਦੇਦਾਰੀ, ਨਿਗਰਾਨੀ ਅਤੇ ਕਾਪੀਰਾਈਟ ਬਾਰੇ ਚਿੰਤਾਵਾਂ ਪੈਦਾ ਕੀਤੀਆਂ ਹਨ।

ਤਕਨਾਲੋਜੀ ਕੰਪਨੀਆਂ ਆਪਣੇ ਉਪਭੋਗਤਾਵਾਂ (ਵਰਤਣ ਵਾਲਿਆਂ) ਤੋਂ ਡਾਟਾ (ਭਿੰਨ ਤਰ੍ਹਾਂ ਦੀ ਜਾਣਕਾਰੀ) ਦੀ ਇਕ ਵਿਸ਼ਾਲ ਲੜੀ ਇਕੱਤਰ ਕਰਦੀਆਂ ਹਨ, ਜਿਸ ਵਿਚ ਆਨਲਾਈਨ ਗਤੀਵਿਧੀ, ਭੂ-ਸਥਾਨ ਜਾਣਕਾਰੀ, ਵੀਡੀਓ ਅਤੇ ਆਡੀਓ ਸ਼ਾਮਲ ਹਨ। ਉਦਾਹਰਣ ਲਈ, ਭਾਸ਼ਣ ਪਛਾਣ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਬਣਾਉਣ ਲਈ, ਐਮਾਜ਼ਾਨ ਹੋਰਾਂ ਨੇ ਲੱਖਾਂ ਨਿਜੀ ਗੱਲਬਾਤਾਂ ਰਿਕਾਰਡ ਕੀਤੀਆਂ ਅਤੇ ਟਾਈਪ ਕਰਨ ਵਾਲਿਆਂ ਨੂੰ ਉਨ੍ਹਾਂ ਵਿਚੋਂ ਕੁਝ ਨੂੰ ਸੁਣਨ ਅਤੇ ਟ੍ਰਾਂਸਕ੍ਰਾਈਬ (ਕੰਪਿਊਟਰ ਵਿਚ ਟਾਈਪ)

ਕਰਨ ਦੀ ਆਗਿਆ ਦਿੱਤੀ। ਇਸ ਵਿਆਪਕ ਨਿਗਰਾਨੀ ਬਾਰੇ ਵਿਚਾਰ ਉਨ੍ਹਾਂ ਲੋਕਾਂ ਤੋਂ ਲੈ ਕੇ ਉਨ੍ਹਾਂ ਲੋਕਾਂ ਤਕ ਹਨ ਜੋ ਇਸ ਨੂੰ ਜ਼ਰੂਰੀ ਬੁਰਾਈ ਵਜੋਂ ਵੇਖਦੇ ਹਨ ਅਤੇ ਜਿਨ੍ਹਾਂ ਲਈ ਇਹ ਸਪੱਸ਼ਟ ਤੌਰ 'ਤੇ ਦੁਰਾਚਾਰੀ ਹੈ ਅਤੇ ਨਿੱਜਤਾ ਦੇ ਅਧਿਕਾਰ ਦੀ ਉਲੰਘਣਾ ਹੈ।

AI ਡਿਵੈਲਪਰਾਂ ਦੀ ਦਲੀਲ ਹੈ ਕਿ ਕੀਮਤੀ ਐਪਲੀਕੇਸ਼ਨਾਂ ਪ੍ਰਦਾਨ ਕਰਨ ਦਾ ਇਹ ਇਕੋ ਇਕ ਤਰੀਕਾ ਹੈ ਅਤੇ ਕਈ ਤਕਨੀਕਾਂ ਵਿਕਸਿਤ ਕੀਤੀਆਂ ਹਨ ਜੋ ਅਜੇ ਵੀ ਡਾਟਾ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) ਪ੍ਰਾਪਤ ਕਰਦੇ ਸਮੇਂ ਪਰਦੇਦਾਰੀ ਨੂੰ ਸੁਰੱਖਿਅਤ ਰੱਖਣ ਦੀ ਕੋਸ਼ਿਸ਼ ਕਰਦੀਆਂ ਹਨ। ਜਿਵੇਂ ਕਿ ਡਾਟਾ ਦੀ ਇਕੱਤਰਤਾ, ਡੀ-ਆਈਡੈਂਟੀਫਿਕੇਸ਼ਨ (ਨਿਜੀ ਜਾਣਕਾਰੀ ਨੂੰ ਲਾਂਭੇ ਕਰਨਾ) ਅਤੇ ਵੱਖਰੀ ਪਰਦੇਦਾਰੀ। 2016 ਤੋਂ, ਕੁਝ ਪਰਦੇਦਾਰੀ ਮਾਹਿਰਾਂ, ਜਿਵੇਂ ਕਿ ਸਿੰਥੀਆ ਡਵਰਕ, ਨੇ ਪਰਦੇਦਾਰੀ ਨੂੰ ਨਿਰਪੱਖਤਾ ਦੇ ਰੂਪ ਵਿਚ ਵੇਖਣਾ ਸ਼ੁਰੂ ਕੀਤਾ - ਬ੍ਰਾਇਨ ਕ੍ਰਿਸਚੀਅਨ ਨੇ ਲਿਖਿਆ ਕਿ ਮਾਹਰ 'ਉਹ ਕੀ ਜਾਣਦੇ ਹਨ' ਦੇ ਸਵਾਲ ਤੋਂ 'ਉਹ ਇਸ ਨਾਲ ਕੀ ਕਰ ਰਹੇ ਹਨ' ਦੇ ਸਵਾਲ ਵੱਲ ਘੁੰਮ ਰਹੇ ਹਨ।

ਜਨਰੇਟਿਵ (ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕੰਪਿਊਟਰ ਸਮੱਗਰੀ ਬਣਾਉਣ ਵਾਲੀ) AI ਨੂੰ ਅਕਸਰ ਗੈਰ-ਲਾਇਸੰਸਸ਼ੁਦਾ ਕਾਪੀਰਾਈਟ ਕੀਤੇ ਕੰਮਾਂ 'ਤੇ ਸਿਖਲਾਈ ਦਿੱਤੀ ਜਾਂਦੀ ਹੈ, ਜਿਸ ਵਿਚ ਚਿੱਤਰ ਜਾਂ ਕੰਪਿਊਟਰ ਕੋਡ ਵਰਗੇ ਕਾਰਜ ਖੇਤਰ ਸ਼ਾਮਲ ਹੁੰਦੇ ਹਨ; ਆਉਟਪੁੱਟ ਨੂੰ ਫਿਰ 'ਵਾਜਬ ਵਰਤੋਂ' ਦੇ ਤਰਕ ਦੇ ਤਹਿਤ ਵਰਤਿਆ ਜਾਂਦਾ ਹੈ। ਮਾਹਰ ਇਸ ਬਾਰੇ ਅਸਹਿਮਤ ਹਨ ਕਿ ਇਹ ਤਰਕ ਕਿੰਨੀ ਚੰਗੀ ਤਰ੍ਹਾਂ, ਅਤੇ ਕਿਹੜੇ ਹਾਲਾਤਾਂ ਵਿਚ, ਕਾਨੂੰਨ ਦੀਆਂ ਅਦਾਲਤਾਂ ਵਿਚ ਕਾਇਮ ਰਹੇਗਾ; ਸਬੰਧਿਤ ਕਾਰਕਾਂ ਵਿਚ "ਕਾਪੀਰਾਈਟ ਕੀਤੇ ਕੰਮ ਦੀ ਵਰਤੋਂ ਦਾ ਉਦੇਸ਼ ਅਤੇ ਚਰਿੱਤਰ" ਅਤੇ "ਕਾਪੀਰਾਈਟ ਕੀਤੇ ਕੰਮ ਲਈ ਸੰਭਾਵਿਤ ਬਾਜ਼ਾਰ ਅਤੇ ਪ੍ਰਭਾਵ" ਸ਼ਾਮਲ ਹੋ ਸਕਦੇ ਹਨ। 2023 ਵਿਚ, ਪ੍ਰਮੁੱਖ ਲੇਖਕਾਂ (ਜੌਨ ਗ੍ਰੀਸ਼ਮ ਅਤੇ ਜੋਨਾਥਨ ਫ੍ਰਾਂਜ਼ਨ ਸਮੇਤ) ਨੇ AI ਕੰਪਨੀਆਂ 'ਤੇ ਜਨਰੇਟਿਵ (ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕੰਪਿਊਟਰ ਸਮੱਗਰੀ ਬਣਾਉਣ ਵਾਲੀ) AI ਨੂੰ ਸਿਖਲਾਈ ਦੇਣ ਲਈ ਆਪਣੇ ਕੰਮ ਦੀ ਵਰਤੋਂ ਕਰਨ ਲਈ ਮੁਕੱਦਮਾ ਦਾਇਰ ਕੀਤਾ।

2. ਗਲਤ ਜਾਣਕਾਰੀ ਤੇ ਪੱਖਪਾਤ

ਯੂ.ਟਿਊਬ, ਫੇਸਬੁੱਕ ਅਤੇ ਹੋਰ ਸੋਸ਼ਲ-ਮੀਡਿਆ ਦਰਸ਼ਕ ਵਧਾਉਣ ਲਈ ਸਿਫ਼ਾਰਿਸ਼ ਪ੍ਰਣਾਲੀਆਂ ਦੀ ਵਰਤੋਂ ਕਰਦੇ ਹਨ। ਇਨ੍ਹਾਂ AI ਪ੍ਰੋਗਰਾਮਾਂ ਨੂੰ ਸਿਫ਼ਾਰਿਸ਼ ਪ੍ਰਣਾਲੀਆਂ ਦੀ ਵਰਤੋਂ ਨੂੰ ਵੱਧ ਤੋਂ ਵੱਧ ਕਰਨ ਦਾ ਟੀਚਾ ਦਿੱਤਾ ਗਿਆ ਸੀ (ਭਾਵ, ਇਕੋ ਇਕ ਟੀਚਾ ਲੋਕਾਂ ਦੇ ਧਿਆਨ ਨੂੰ ਆਕਰਸ਼ਕ ਕਰਨਾ ਸੀ)। ਇਸ ਤਰ੍ਹਾਂ AI ਨੇ ਸਿੱਖਿਆ ਕਿ ਲੋਕਾਂ ਦੇ ਧਿਆਨ ਨੂੰ ਆਕਰਸ਼ਕ ਕਰਨ ਵਾਸਤੇ ਗਲਤ ਜਾਣਕਾਰੀ, ਸਾਜ਼ਿਸ਼ ਦੇ ਸਿਧਾਂਤਾਂ ਅਤੇ ਅਤਿਅੰਤ ਪੱਖਪਾਤੀ ਸਮੱਗਰੀ ਦੀ ਚੋਣ ਲਾਭਦਾਇਕ ਹੈ, ਅਤੇ, ਉਨ੍ਹਾਂ ਦੇ ਵੇਖਦੇ ਰਹਿਣ ਲਈ, AI ਦੀ ਵਧੇਰੇ ਸਿਫ਼ਾਰਿਸ਼ ਕੀਤੀ ਜਾਂਦੀ ਹੈ। ਐਸੀ ਵਿਧੀਆਂ ਨੇ ਉਸੇ ਵਿਸ਼ੇ 'ਤੇ ਵਧੇਰੇ ਸਮੱਗਰੀ ਵੇਖਣ ਦਾ ਰੁਝਾਨ ਵੀ ਰੱਖਿਆ, ਇਸ ਲਈ AI ਨੇ ਲੋਕਾਂ ਨੂੰ ਖਾਸ ਘੇਰੇ ਵਿਚ ਲਿਜਾਇਆ ਜਿਥੇ ਉਨ੍ਹਾਂ ਨੂੰ ਇੱਕੋ ਗਲਤ ਜਾਣਕਾਰੀ ਦੇ ਕਈ ਸੰਸਕਰਣ ਪ੍ਰਾਪਤ ਹੋਏ। ਇਸ ਨੇ ਬਹੁਤ ਸਾਰੇ ਉਪਭੋਗਤਾਵਾਂ (ਵਰਤੋਂ ਕਰਨ ਵਾਲਿਆਂ) ਨੂੰ ਯਕੀਨ ਦਿਵਾਇਆ ਕਿ ਗਲਤ ਜਾਣਕਾਰੀ ਸੱਚ ਸੀ, ਅਤੇ ਆਖਰਕਾਰ ਸੰਸਥਾਵਾਂ, ਮੀਡਿਆ ਅਤੇ ਸਰਕਾਰ ਵਿਚ ਵਿਸ਼ਵਾਸ ਨੂੰ ਕਮਜ਼ੋਰ ਕਰ ਦਿੱਤਾ। AI ਪ੍ਰੋਗਰਾਮ ਨੇ ਆਪਣੇ ਟੀਚੇ ਨੂੰ ਵੱਧ ਤੋਂ ਵੱਧ ਕਰਨਾ ਸਹੀ ਢੰਗ ਨਾਲ ਸਿੱਖ ਲਿਆ ਸੀ, ਪਰ ਨਤੀਜਾ ਸਮਾਜ ਲਈ ਨੁਕਸਾਨ-ਦੇਹ ਸੀ। ਸਾਲ 2016 'ਚ ਅਮਰੀਕੀ ਚੋਣਾਂ ਤੋਂ ਬਾਅਦ ਵੱਡੀਆਂ ਤਕਨਾਲੋਜੀ ਕੰਪਨੀਆਂ ਨੇ ਇਸ ਸਮੱਸਿਆ ਨੂੰ ਘੱਟ ਕਰਨ ਲਈ ਕਦਮ ਚੁੱਕੇ ਸਨ। ਪਰ ਉਸ ਦੇ ਉਲਟ ਕਈਆਂ ਨੇ ਇਸ ਪੱਖਪਾਤ ਨੂੰ ਵਧਾਉਣ ਲਈ ਕਦਮ ਚੁੱਕੇ ਹਨ।

2022 ਵਿਚ, ਜਨਰੇਟਿਵ (ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕੰਪਿਊਟਰ ਸਮੱਗਰੀ ਬਣਾਉਣ ਵਾਲੀ) AI ਨੇ ਚਿੱਤਰ, ਆਡੀਓ, ਵੀਡੀਓ ਅਤੇ ਟੈੱਕਸਟ ਬਣਾਉਣਾ ਸ਼ੁਰੂ ਕੀਤਾ ਜੋ ਅਸਲ ਫੋਟੋਆਂ, ਰਿਕਾਰਡਿੰਗਾਂ, ਫਿਲਮਾਂ ਜਾਂ ਮਨੁੱਖੀ ਲਿਖਤ ਤੋਂ ਵੱਖਰੇ ਨਹੀਂ ਹਨ। ਮਾੜੇ ਅਨਸਰਾਂ ਲਈ ਵੱਡੀ ਮਾਤਰਾ ਵਿਚ ਗਲਤ ਜਾਣਕਾਰੀ ਜਾਂ ਪ੍ਰਚਾਰ ਪੈਦਾ ਕਰਨ ਲਈ ਇਸ ਤਕਨਾਲੋਜੀ ਦੀ ਵਰਤੋਂ ਕਰਨਾ ਸੰਭਵ ਹੈ। ਇਸ ਤਕਨਾਲੋਜੀ ਨੂੰ ਘੱਟੋ ਘੱਟ ਲਾਗਤ 'ਤੇ ਵਿਆਪਕ ਤੌਰ 'ਤੇ ਵੰਡਿਆ ਗਿਆ ਹੈ। ਹੁਣ ਇਹ ਹਾਲਤ ਹੈ ਕਿ ਕੋਈ ਵੀ ਬਨਾਵਟੀ ਵੀਡੀਓ, ਆਡੀਓ ਜਾਂ ਫ਼ੋਟੋਆਂ ਬਣਾ ਸਕਦਾ ਹੈ ਜਿਨ੍ਹਾਂ ਨੂੰ ਅਸਲ ਨਾਲੋਂ ਵੱਖ ਦੱਸਣਾ ਸੰਭਵ ਨਹੀਂ। ਇਸ ਤਰ੍ਹਾਂ ਗਲਤ ਅਫਵਾਹਾਂ/ਖ਼ਬਰਾਂ ਦਾ ਫ਼ੈਲਾਉਣਾ ਬਹੁਤ ਸੌਖਾ ਹੋ ਗਿਆ ਹੈ ਅਤੇ ਇਸ ਤਰ੍ਹਾਂ ਸੱਚ ਤੇ ਝੂਠ ਦਾ ਫਰਕ ਦੱਸਣਾ ਔਖਾ ਹੋ ਗਿਆ ਹੈ। ਇਸ ਪ੍ਰਗਤੀ ਦੇ ਬਹੁਤ ਗੰਭੀਰ ਨਤੀਜੇ ਹੋਣਗੇ। ਜਿਓਫਰੀ ਹਿੰਟਨ (ਜੋ ਇਨ੍ਹਾਂ ਸਾਧਨਾਂ ਦਾ

ਇਕ ਮਹੱਤਵਪੂਰਣ ਵਿਕਾਸ-ਕਰਤਾ ਸੀ) ਨੇ AI ਦੀ ਗਲਤ ਜਾਣਕਾਰੀ ਬਾਰੇ ਆਪਣੀਆਂ ਚਿੰਤਾਵਾਂ ਜ਼ਾਹਰ ਕੀਤੀਆਂ। ਉਸਨੇ AI ਵਿਕਸਤ ਕਰਨ ਵਾਲੀਆਂ ਕੰਪਨੀਆਂ ਦੀ ਖੁੱਲ੍ਹ ਕੇ ਆਲੋਚਨਾ ਕਰਨ ਲਈ ਅਤੇ AI ਦੇ ਖ਼ਤਰੇ ਦੱਸਣ ਲਈ ਗੂਗਲ ਦੀ ਨੌਕਰੀ ਛੱਡ ਦਿੱਤੀ।

ਉਪਰ ਦੱਸੇ ਗਏ ਤਰੀਕੇ ਬਹੁਤ ਸਾਰੇ ਸਿਆਸੀ ਅਤੇ ਹੋਰ ਵਰਗਾਂ ਦੇ ਆਗੂਆਂ ਅਤੇ ਹੋਰਾਂ ਲਈ ਵਧੀਆ ਢੰਗ (ਟੂਲ) ਬਣ ਗਏ ਹਨ ਕਿ ਲੋਕਾਂ ਨੂੰ ਕਿਵੇਂ ਪ੍ਰਭਾਵਿਤ ਕਰਨਾ ਹੈ। AI ਦੀ ਵਰਤੋਂ ਤੋਂ ਪਹਿਲਾਂ ਹੀ ਸਮਾਜਕ/ਸੋਸ਼ਲ-ਮੀਡਿਆ ਦੀ ਵਿਸ਼ਾਲ ਪਹੁੰਚ ਕਾਰਨ ਝੂਠ ਤੇ ਪ੍ਰਾਪੇਗੰਡਾ ਫੈਲਾਉਣਾ ਸੌਖਾ ਹੋ ਗਿਆ ਸੀ, ਪਰ AI ਦੇ ਐਲਗੋਰਿਦਮਾ ਨਾਲ ਹੋਰ ਵੀ ਸੌਖਾ ਹੋ ਗਿਆ ਹੈ। ਸੋਸ਼ਲ-ਮੀਡਿਆ ਨੇ ਜੋ ਸਮਾਜਕ ਦਰਾੜਾਂ ਪਾਈਆਂ ਸੀ ਉਹ ਹੋਰ ਡੂੰਘੀਆਂ ਹੋ ਰਹੀਆਂ ਹਨ।

3. **ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਕਾਰਨ ਪੱਖਪਾਤ ਅਤੇ ਨਿਰਪੱਖਤਾ**

ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦਾ ਪੱਖਪਾਤ ਕੰਪਿਊਟਰ ਸਿਸਟਮ ਵਿਚ ਯੋਜਨਾਬੱਧ ਅਤੇ ਦੁਹਰਾਉਣ-ਯੋਗ ਗਲਤੀਆਂ ਦਾ ਵਰਣਨ ਕਰਦਾ ਹੈ ਜੋ "ਅਨੁਚਿਤ/ਇਤਰਾਜ਼-ਯੋਗ" ਨਤੀਜੇ ਪੈਦਾ ਕਰਦੇ ਹਨ, ਜਿਵੇਂ ਕਿ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੇ ਇਰਾਦੇ ਵਾਲੇ ਕਾਰਜ ਤੋਂ ਵੱਖਰੇ ਤਰੀਕਿਆਂ ਨਾਲ ਇਕ ਸ਼੍ਰੇਣੀ ਨੂੰ ਦੂਜੀ ਸ਼੍ਰੇਣੀ 'ਤੇ "ਨਿੱਜੀਕਰਨ" ਕਰਨਾ। ਪੱਖਪਾਤ ਬਹੁਤ ਸਾਰੇ ਕਾਰਨਾਂ ਤੋਂ ਉਭਰ ਸਕਦਾ ਹੈ, ਜਿਸ ਵਿਚ ਐਲਗੋਰਿਦਮ ਨੂੰ ਸਿਖਲਾਈ ਦੇਣ ਲਈ ਡੇਟਾ (ਜਾਣਕਾਰੀ) ਨੂੰ ਕੋਡ ਕਰਨ, ਇਕੱਤਰ ਕਰਨ, ਚੁਣਨ ਜਾਂ ਵਰਤਣ ਦੇ ਤਰੀਕੇ ਨਾਲ ਸਬੰਧਿਤ ਅਣਚਾਹੇ ਜਾਂ ਅਣਕਿਆਸੇ ਵਰਤੋਂ ਜਾਂ ਫੈਸਲੇ ਸ਼ਾਮਲ ਹਨ ਪਰ ਇਥੋਂ ਤਕ ਸੀਮਤ ਨਹੀਂ ਹਨ।

ਉਦਾਹਰਣ ਲਈ, ਆਧੁਨਿਕ ਖੋਜ ਇੰਜਣ ਨਤੀਜੇ ਅਤੇ ਸੋਸ਼ਲ-ਮੀਡਿਆ ਪਲੈਟਫਾਰਮ (ਉਤਰ ਸਥਿਤੀ) ਵਿਚ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦਾ ਪੱਖਪਾਤ ਦੇਖਿਆ ਗਿਆ ਹੈ। ਇਸ ਪੱਖਪਾਤ ਨਾਲ ਅਣਜਾਣੇ ਵਿਚ ਸਮਾਜਕ ਪੱਖਪਾਤ, ਨਸਲ, ਲਿੰਗ, ਲਿੰਗਕਤਾ, ਅਤੇ ਨਸਲੀ ਵਿਤਕਰੇ ਹੋ ਸਕਦੇ ਹਨ। ਇਹ ਇਕ ਗੰਭੀਰ ਵਿਸ਼ਾ ਹੈ ਜਿਸ ਦਾ ਹਲ ਸੌਖਾ ਨਹੀਂ ਹੈ।

ਮਸ਼ੀਨ-ਲਰਨਿੰਗ (ਸਿਖਲਾਈ) ਐਪਲੀਕੇਸ਼ਨਾਂ ਪੱਖਪਾਤੀ ਹੋਣਗੀਆਂ ਜੇ ਉਹ ਪੱਖਪਾਤੀ ਡਾਟਾ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) ਤੋਂ ਸਿੱਖਦੀਆਂ ਹਨ। ਪਰ ਡਿਵੈਲਪਰਾਂ ਨੂੰ ਪਤਾ ਨਹੀਂ ਹੋ ਸਕਦਾ ਹੈ ਕਿ ਪੱਖਪਾਤ ਮੌਜੂਦ ਹੈ। ਪੱਖਪਾਤ ਨੂੰ ਸਿਖਲਾਈ ਡਾਟਾ (ਜਾਣਕਾਰੀ) ਦੀ ਚੋਣ ਕਰਨ ਦੇ ਤਰੀਕੇ ਅਤੇ ਇਕ ਮਾਡਲ ਨੂੰ ਤਾਇਨਾਤ ਕਰਨ ਦੇ ਤਰੀਕੇ ਦੁਆਰਾ ਪੇਸ਼ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ। ਜੇ ਕਿਸੇ ਪੱਖਪਾਤੀ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੀ ਵਰਤੋਂ ਅਜਿਹੇ ਫੈਸਲੇ ਲੈਣ ਲਈ ਕੀਤੀ ਜਾਂਦੀ ਹੈ ਜੋ ਲੋਕਾਂ ਨੂੰ ਗੰਭੀਰ ਤੌਰ 'ਤੇ ਨੁਕਸਾਨ ਪਹੁੰਚਾ ਸਕਦੇ ਹਨ (ਜਿਵੇਂ ਕਿ ਇਹ ਨੁਕਸਾਨ ਦਵਾਈ, ਵਿੱਤ, ਭਰਤੀ, ਰਿਹਾਇਸ਼ ਜਾਂ ਪੁਲਿਸ ਦੇ ਕੰਮ ਵਿਚ ਹੋ ਸਕਦਾ ਹੈ) ਤਾਂ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਵਿਚ ਭੇਦ ਭਾਵ ਦਾ ਕਾਰਨ ਬਣ ਸਕਦਾ ਹੈ। ਮਸ਼ੀਨ-ਲਰਨਿੰਗ (ਸਿਖਲਾਈ) ਵਿਚ ਨਿਰਪੱਖਤਾ ਇਸ ਗੱਲ ਦਾ ਅਧਿਐਨ ਹੈ ਕਿ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੇ ਪੱਖਪਾਤ ਕਾਰਨ ਹੋਣ ਵਾਲੇ ਨੁਕਸਾਨ ਨੂੰ ਕਿਵੇਂ ਰੋਕਿਆ ਜਾਵੇ। ਇਹ AI ਦੇ ਅੰਦਰ ਅਕਾਦਮਿਕ ਅਧਿਐਨ ਦਾ ਗੰਭੀਰ ਖੇਤਰ ਬਣ ਗਿਆ ਹੈ। ਖੋਜ-ਕਰਤਾ ਅਧਿਕਾਰੀਆਂ ਨੇ ਖੋਜ ਕੀਤੀ ਹੈ ਕਿ "ਨਿਰਪੱਖਤਾ" ਨੂੰ ਇਸ ਤਰੀਕੇ ਨਾਲ ਪਰਿਭਾਸ਼ਿਤ ਕਰਨਾ ਹਮੇਸ਼ਾ ਸੰਭਵ ਨਹੀਂ ਹੁੰਦਾ ਜੋ ਸਾਰੇ ਹਿੱਸੇਦਾਰਾਂ ਨੂੰ ਸੰਤੁਸ਼ਟ ਕਰਦਾ ਹੋਵੇ।

28 ਜੂਨ, 2015 ਨੂੰ, ਗੂਗਲ-ਫੋਟੋਜ਼ ਦੇ ਨਵੇਂ ਚਿੱਤਰ-ਲੇਬਲਿੰਗ-ਫੀਚਰ ਨੇ ਗਲਤੀ ਨਾਲ ਜੈਕੀ ਅਲਸੀਨ ਅਤੇ ਇਕ ਦੋਸਤ ਨੂੰ "ਗੋਰਿਲਾ" ਵਜੋਂ ਪਛਾਣ ਲਿਆ ਕਿਉਂਕਿ ਉਹ ਕਾਲੇ ਸਨ। ਸਿਸਟਮ ਨੂੰ ਇਕ ਡਾਟਾ-ਸੈਟ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) 'ਤੇ ਸਿਖਲਾਈ ਦਿੱਤੀ ਗਈ ਸੀ ਜਿਸ ਵਿਚ ਕਾਲੇ ਲੋਕਾਂ ਦੀਆਂ ਬਹੁਤ ਘੱਟ ਤਸਵੀਰਾਂ ਸਨ, ਇਕ ਸਮੱਸਿਆ ਜਿਸ ਨੂੰ "ਨਮੂਨੇ ਦੇ ਆਕਾਰ ਦੀ ਅਸਮਾਨਤਾ" ਕਿਹਾ ਜਾਂਦਾ ਹੈ। ਗੂਗਲ ਨੇ ਸਿਸਟਮ ਨੂੰ ਕਿਸੇ ਵੀ ਚੀਜ਼ ਨੂੰ "ਗੋਰਿਲਾ" ਵਜੋਂ ਲੇਬਲ ਕਰਨ ਤੋਂ ਰੋਕ ਕੇ ਇਸ ਸਮੱਸਿਆ ਨੂੰ ਠੀਕ ਕੀਤਾ ਹੈ। ਇਸ ਕਾਰਨ, ਅੱਠ ਸਾਲ ਬਾਅਦ, 2023 ਵਿਚ, ਗੂਗਲ-ਫੋਟੋਜ਼ ਅਜੇ ਵੀ ਗੋਰਿਲਾ ਦੀ ਪਛਾਣ ਨਹੀਂ ਕਰ ਸਕੀ, ਅਤੇ ਨਾ ਹੀ ਐਪਲ, ਫੇਸਬੁੱਕ, ਮਾਈਕ੍ਰੋਸਾਫਟ ਅਤੇ ਐਮਾਜ਼ਾਨ ਦੇ ਸਮਾਨ ਉਤਪਾਦਾਂ ਦੀ ਪਛਾਣ ਕਰ ਸਕੀ।

ਕੰਪਸ (COMPAS) ਇਕ ਵਪਾਰਕ ਪ੍ਰੋਗਰਾਮ ਹੈ ਜੋ ਅਮਰੀਕੀ ਅਦਾਲਤਾਂ ਦੁਆਰਾ ਵਿਆਪਕ ਤੌਰ 'ਤੇ ਵਰਤਿਆ ਜਾਂਦਾ ਹੈ ਤਾਂ ਜੋ ਕਿਸੇ ਬਚਾਓ ਕਰਤਾ ਦੇ ਲਾਭਦਾਇਕ ਬਣਨ ਦੀ ਸੰਭਾਵਨਾ ਦਾ ਮੁਲਾਂਕਣ

ਕੀਤਾ ਜਾ ਸਕੇ। 2016 ਵਿਚ, ਪ੍ਰੋਪਬਲਿਕਾ ਵਿਖੇ ਜੂਲੀਆ ਐਂਗਲਵਿਨ ਨੇ ਪਾਇਆ ਕਿ ਕੰਪਸ (COMPAS) ਨੇ ਨਸਲੀ ਪੱਖਪਾਤ ਦਾ ਪ੍ਰਦਰਸ਼ਨ ਕੀਤਾ। ਇਸ ਤੱਥ ਦੇ ਬਾਵਜੂਦ ਕਿ ਪ੍ਰੋਗਰਾਮ ਨੂੰ ਬਚਾਓ ਪਰਨਾਲੀ ਦੀਆਂ ਨਸਲਾਂ ਬਾਰੇ ਨਹੀਂ ਦੱਸਿਆ ਗਿਆ ਸੀ। ਹਾਲਾਂਕਿ ਗੋਰੇ ਅਤੇ ਕਾਲੇ ਦੋਵਾਂ ਲਈ ਗਲਤੀ ਦੀ ਦਰ ਬਿਲਕੁਲ 61% ਦੇ ਬਰਾਬਰ ਸੀ, ਪਰ ਹਰੇਕ ਨਸਲ ਲਈ ਗਲਤੀਆਂ ਵੱਖਰੀਆਂ ਸਨ - ਸਿਸਟਮ ਨੇ ਲਗਾਤਾਰ ਇਸ ਸੰਭਾਵਨਾ ਨੂੰ ਵਧਾ-ਚੜ੍ਹਾ ਕੇ ਪੇਸ਼ ਕੀਤਾ ਕਿ ਇਕ ਕਾਲਾ ਵਿਅਕਤੀ ਦੁਬਾਰਾ ਗ੍ਰਿਫ਼ਤਾਰ ਹੋਵੇਗਾ ਅਤੇ ਇਸ ਸੰਭਾਵਨਾ ਨੂੰ ਘੱਟ ਪੇਸ਼ ਕੀਤਾ ਕਿ ਇਕ ਗੋਰਾ ਵਿਅਕਤੀ ਦੁਬਾਰਾ ਗ੍ਰਿਫ਼ਤਾਰ ਹੋਵੇਗਾ। 2017 ਵਿਚ, ਕਈ ਖੋਜ-ਕਰਤਾ ਅਧਿਕਾਰੀਆਂ ਨੇ ਦਿਖਾਇਆ ਕਿ ਕੰਪਸ ਲਈ ਨਿਰਪੱਖਤਾ ਦੇ ਸਾਰੇ ਸੰਭਵ ਉਪਾਵਾਂ ਨੂੰ ਸ਼ਾਮਲ ਕਰਨਾ ਗਿਣਤੀ ਅਨੁਸਾਰ ਅਸੰਭਵ ਸੀ ਜਦੋਂ ਕਿ ਅੰਕੜਿਆਂ ਵਿਚ ਗੋਰੇ ਅਤੇ ਕਾਲੇ ਲੋਕਾਂ ਲਈ ਦੁਬਾਰਾ ਅਪਰਾਧ ਦੀਆਂ ਮੂਲ ਦਰਾਂ ਵੱਖਰੀਆਂ ਸਨ।

ਇਕ ਪ੍ਰੋਗਰਾਮ ਪੱਖਪਾਤੀ ਫੈਸਲੇ ਲੈ ਸਕਦਾ ਹੈ ਭਾਵੇਂ ਡਾਟਾ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ ਸੂਚੀ) ਸਪੱਸ਼ਟ ਤੌਰ 'ਤੇ ਕਿਸੇ ਸਮੱਸਿਆ ਵਾਲੀ ਵਿਸ਼ੇਸ਼ਤਾ (ਜਿਵੇਂ ਕਿ "ਨਸਲ" ਜਾਂ "ਲਿੰਗ") ਦਾ ਜ਼ਿਕਰ ਨਹੀਂ ਕਰਦਾ। ਇਹ ਵਿਸ਼ੇਸ਼ਤਾ ਹੋਰ ਵਿਸ਼ੇਸ਼ਤਾਵਾਂ (ਜਿਵੇਂ ਕਿ "ਪਤਾ", "ਖਰੀਦਦਾਰੀ ਦਾ ਇਤਿਹਾਸ" ਜਾਂ "ਪਹਿਲਾ ਨਾਮ") ਨਾਲ ਸਬੰਧਿਤ ਹੋਵੇਗੀ, ਅਤੇ ਪ੍ਰੋਗਰਾਮ ਇਹਨਾਂ ਵਿਸ਼ੇਸ਼ਤਾਵਾਂ ਦੇ ਅਧਾਰ 'ਤੇ ਉਹੀ ਫੈਸਲੇ ਲਵੇਗਾ ਜੋ ਇਹ "ਨਸਲ" ਜਾਂ "ਲਿੰਗ" 'ਤੇ ਕਰੇਗਾ। ਮੋਰਿਟਜ਼ ਹਾਰਡਟ ਨੇ ਕਿਹਾ, "ਇਸ ਖੋਜ ਖੇਤਰ ਵਿਚ ਸਭ ਤੋਂ ਮਜ਼ਬੂਤ ਤੱਥ ਇਹ ਹੈ ਕਿ ਅੰਨ੍ਹੇਪਣ ਰਾਹੀਂ ਨਿਰਪੱਖਤਾ ਕੰਮ ਨਹੀਂ ਹੋ ਕਰਦੀ। ਪੱਖਪਾਤ ਅਤੇ ਬੇਇਨਸਾਫ਼ੀ ਦਾ ਪਤਾ ਇਸ ਕਰਕੇ ਵੀ ਨਹੀਂ ਲੱਗ ਸਕਦਾ ਕਿਉਂਕਿ ਡਿਵੈਲਪਰ ਬਹੁਤ ਜ਼ਿਆਦਾ ਗੋਰੇ ਲੋਕ ਅਤੇ ਪੁਰਸ਼ ਹਨ। AI ਇੰਜੀਨੀਅਰਾਂ ਵਿਚ, ਲਗਭਗ 4% ਕਾਲੇ ਅਤੇ 20% ਔਰਤਾਂ ਹਨ।

ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੇ ਪੱਖਪਾਤ ਦਾ ਅਧਿਐਨ ਐਲਗੋਰਿਦਮ ਨਾਲ ਸਭ ਤੋਂ ਵੱਧ ਸਬੰਧਿਤ ਹੈ ਜੋ "ਲਿੰਗ ਅਤੇ ਨਸਲ" ਵਿਤਕਰੇ ਨੂੰ ਦਰਸਾਉਂਦੇ ਹਨ। ਇਸ ਪੱਖਪਾਤ ਨੂੰ ਹਾਲ ਹੀ ਵਿਚ ਕਾਨੂੰਨੀ ਢਾਂਚੇ ਵਿਚ ਸੰਬੋਧਿਤ ਕੀਤਾ ਗਿਆ ਹੈ, ਜਿਵੇਂ ਕਿ ਯੂਰੋਪੀਅਨ ਯੂਨੀਅਨ ਦੇ ਜਨਰਲ ਡਾਟਾ ਪ੍ਰੋਟੈਕਸ਼ਨ ਰੈਗੂਲੇਸ਼ਨ (2018) ਅਤੇ ਪ੍ਰਸਤਾਵਿਤ ਆਰਟੀਫੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਐਕਟ (2021)।

ਜਿਵੇਂ ਕਿ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਸਮਾਜ, ਰਾਜਨੀਤੀ, ਸੰਸਥਾਵਾਂ ਅਤੇ ਵਿਵਹਾਰ ਨੂੰ ਸੰਗਠਿਤ ਕਰਨ ਦੀ ਆਪਣੀ ਯੋਗਤਾ ਦਾ ਵਿਸਥਾਰ ਕਰਦੇ ਹਨ, ਸਮਾਜ-ਵਿਗਿਆਨੀ ਉਹਨਾਂ ਤਰੀਕਿਆਂ ਨਾਲ ਚਿੰਤਤ ਹੋ ਗਏ ਹਨ ਜਿਨ੍ਹਾਂ ਵਿਚ ਅਣਕਿਆਸੇ ਆਉਟਪੁੱਟ ਅਤੇ ਡੇਟਾ ਦੀ ਹੇਰਾਫੇਰੀ ਭੌਤਿਕ ਸੰਸਾਰ ਨੂੰ ਪ੍ਰਭਾਵਿਤ ਕਰ ਸਕਦੀ ਹੈ। ਪੱਖਪਾਤ ਪਹਿਲਾਂ ਤੋਂ ਮੌਜੂਦ ਸਭਿਆਚਾਰਕ, ਸਮਾਜਕ, ਜਾਂ ਸੰਸਥਾਗਤ ਉਮੀਦਾਂ ਦੇ ਨਤੀਜੇ ਵਜੋਂ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ) ਦੀ ਪ੍ਰਨਾਲੀਆਂ ਵਿਚ ਦਾਖਲ ਹੋ ਸਕਦਾ ਹੈ; ਉਹਨਾਂ ਦੇ ਡਿਜ਼ਾਈਨ (ਵਿਓਂਤ) ਦੀਆਂ ਤਕਨੀਕੀ ਕਮੀਆਂ ਦੇ ਕਾਰਨ; ਜਾਂ ਅਣਪਛਾਤੇ ਸੰਦਰਭਾਂ ਵਿਚ ਜਾਂ ਉਹਨਾਂ ਦਰਸ਼ਕਾਂ ਦੁਆਰਾ ਵਰਤੇ ਜਾਣ ਦੁਆਰਾ ਜਿਨ੍ਹਾਂ ਨੂੰ ਸਾਫਟਵੇਅਰ ਦੇ ਸ਼ੁਰੂ ਵਾਲੀ ਡਿਜ਼ਾਈਨ (ਵਿਓਂਤ) ਵਿਚ ਨਹੀਂ ਮੰਨਿਆ ਜਾਂਦਾ ਹੈ।

ਚੋਣ ਨਤੀਜਿਆਂ ਤੋਂ ਲੈ ਕੇ ਫੈਲਣ ਤਕ ਦੇ ਮਾਮਲਿਆਂ ਵਿਚ ਐਲਗੋਰਿਦਮ ਦੇ ਪੱਖਪਾਤ ਦਾ ਹਵਾਲਾ ਦਿੱਤਾ ਗਿਆ ਹੈ ਜਿਵੇਂ ਕਿ ਆਨਲਾਈਨ ਨਫ਼ਰਤ ਭਾਸ਼ਣ। ਇਹ ਅਪਰਾਧਿਕ ਨਿਆਂ, ਸਿਹਤ ਸੰਭਾਲ, ਮੌਜੂਦਾ ਨਸਲੀ, ਸਮਾਜਕ-ਆਰਥਿਕ, ਅਤੇ ਲਿੰਗ ਪੱਖਪਾਤ ਨੂੰ ਮਿਸ਼ਰਿਤ ਕਰਨ ਵਿਚ ਵੀ ਪੈਦਾ ਹੋਇਆ ਹੈ। ਗੂੜ੍ਹੀ ਚਮੜੀ ਵਾਲੇ ਚਿਹਰਿਆਂ ਦੀ ਸਹੀ ਪਛਾਣ ਕਰਨ ਲਈ ਚਿਹਰੇ ਦੀ ਪਛਾਣ ਕਰਨ ਵਾਲੀ ਤਕਨਾਲੋਜੀ ਦੀ ਬਣਤਰ ਅਸਮਰੱਥਾ ਨੂੰ ਕਾਲੇ ਆਦਮੀਆਂ ਦੀਆਂ ਕਈ ਗਲਤ ਗ੍ਰਿਫ਼ਤਾਰੀਆਂ ਨਾਲ ਜੋੜਿਆ ਗਿਆ ਹੈ, ਇਕ ਮੁੱਦਾ ਅਸੰਤੁਲਿਤ ਡੇਟਾਸੈਟਾਂ (ਇਕੱਠੀ ਕੀਤੀ ਜਾਣਕਾਰੀ) ਤੋਂ ਪੈਦਾ ਹੁੰਦਾ ਹੈ। ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੀ ਮਲਕੀਅਤ ਦੇ ਕਾਰਨ ਐਲਗੋਰਿਦਮ ਦੇ ਪੱਖਪਾਤ ਨੂੰ ਸਮਝਣ, ਖੋਜ ਕਰਨ ਅਤੇ ਖੋਜਣ ਵਿਚ ਸਮੱਸਿਆਵਾਂ ਬਰਕਰਾਰ ਰਹਿੰਦੀਆਂ ਹਨ, ਜਿਨ੍ਹਾਂ ਨੂੰ ਆਮ ਤੌਰ 'ਤੇ ਵਪਾਰਕ ਰਾਜ਼ ਮੰਨਿਆ ਜਾਂਦਾ ਹੈ। ਇਥੋਂ ਤਕ ਕਿ ਜਦੋਂ ਪੂਰੀ ਪਾਰਦਰਸ਼ਤਾ ਪ੍ਰਦਾਨ ਕੀਤੀ ਜਾਂਦੀ ਹੈ, ਕੁਝ ਐਲਗੋਰਿਦਮ ਦੀ ਗੁੰਝਲਤਾ ਉਹਨਾਂ ਦੇ ਕੰਮਕਾਜ ਨੂੰ ਸਮਝਣ ਵਿਚ ਰੁਕਾਵਟ ਬਣਾਉਂਦੀ ਹੈ। ਇਸ ਤੋਂ ਇਲਾਵਾ, ਐਲਗੋਰਿਦਮ ਬਦਲ ਸਕਦੇ ਹਨ, ਜਾਂ ਇਨਪੁੱਟ ਜਾਂ ਆਉਟਪੁੱਟ ਨੂੰ ਅਜਿਹੇ ਤਰੀਕਿਆਂ ਨਾਲ ਜਵਾਬ ਦੇ ਸਕਦੇ ਹਨ ਜਿਨ੍ਹਾਂ ਦਾ ਅਨੁਮਾਨ ਨਹੀਂ ਲਗਾਇਆ ਜਾ

ਸਕਦਾ ਜਾਂ ਵਿਸ਼ਲੇਸ਼ਣ ਲਈ ਆਸਾਨੀ ਨਾਲ ਦੁਬਾਰਾ ਤਿਆਰ ਨਹੀਂ ਕੀਤਾ ਜਾ ਸਕਦਾ। ਬਹੁਤ ਸਾਰੇ ਮਾਮਲਿਆਂ ਵਿਚ, ਇਥੋਂ ਤਕ ਕਿ ਇਕ ਵੈਬਸਾਈਟ ਜਾਂ ਐਪਲੀਕੇਸ਼ਨ ਦੇ ਅੰਦਰ, ਜਾਂਚ ਕਰਨ ਲਈ ਕੋਈ ਇਕੱਲਾ "ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ)" ਨਹੀਂ ਹੈ, ਪਰ ਕਈ ਆਪਸ ਵਿਚ ਜੁੜੇ ਪ੍ਰੋਗਰਾਮਾਂ ਅਤੇ ਡੇਟਾ ਇਨਪੁੱਟ ਦਾ ਇਕ ਨੈੱਟਵਰਕ, ਇਥੋਂ ਤਕ ਕਿ ਇੱਕੋ ਸੇਵਾ ਦੇ ਉਪਭੋਗਤਾਵਾਂ (users) ਵਿਚਕਾਰ ਵੀ ਨਹੀਂ।

4. ਸਪੱਸ਼ਟਤਾ ਦੀ ਘਾਟ

ਜ਼ਿਆਦਾਤਰ ਆਧੁਨਿਕ AI ਐਪਲੀਕੇਸ਼ਨਾਂ ਇਹ ਨਹੀਂ ਦੱਸ ਸਕਦੀਆਂ ਕਿ ਉਹ ਕਿਸੇ ਫੈਸਲੇ 'ਤੇ ਕਿਵੇਂ ਪਹੁੰਚੇ ਹਨ। ਡੂੰਘੇ ਨਿਊਰਲ ਨੈੱਟਵਰਕ ਵਿਚ ਇਨਪੁਟਾਂ ਅਤੇ ਆਉਟਪੁੱਟਾਂ ਵਿਚਕਾਰ ਵੱਡੀ ਮਾਤਰਾ ਵਿਚ ਸਬੰਧ ਅਤੇ ਨਤੀਜੇ ਵਜੋਂ ਦੀ ਗੁੰਝਲ ਨੂੰ ਇਕ ਮਾਹਰ ਲਈ ਵੀ ਇਹ ਸਮਝਾਉਣਾ ਮੁਸ਼ਕਲ ਬਣਾਉਂਦੀ ਹੈ ਕਿ ਉਨ੍ਹਾਂ ਨੇ ਆਪਣੇ ਆਉਟਪੁੱਟ ਕਿਵੇਂ ਪੈਦਾ ਕੀਤੇ, ਜਿਸ ਨਾਲ ਉਹ ਇਕ ਬਲੈਕ-ਬਾਕਸ (ਗੁੰਝਲਦਾਰ ਡੱਬਾ) ਬਣ ਗਏ।

ਅਜਿਹੇ ਬਹੁਤ ਸਾਰੇ ਮਾਮਲੇ ਸਾਹਮਣੇ ਆਏ ਹਨ ਜਿਥੇ ਇਕ ਮਸ਼ੀਨ-ਲਰਨਿੰਗ ਪ੍ਰੋਗਰਾਮ ਨੇ ਸਖ਼ਤ ਟੈਸਟ ਪਾਸ ਕੀਤੇ ਹਨ, ਪਰ ਫਿਰ ਵੀ ਪ੍ਰੋਗਰਾਮਰਾਂ ਦੇ ਇਰਾਦੇ ਨਾਲੋਂ ਕੁਝ ਵੱਖਰਾ ਸਿੱਖਿਆ ਹੈ। ਉਦਾਹਰਣ ਵਜੋਂ, ਜਸਟਿਨ ਕੋ ਅਤੇ ਰਾਬਰਟੋ ਨੋਵੋਆ ਨੇ ਇਕ ਪ੍ਰਣਾਲੀ ਵਿਕਸਿਤ ਕੀਤੀ ਜੋ ਡਾਕਟਰੀ ਪੇਸ਼ੇਵਰਾਂ ਨਾਲੋਂ ਚਮੜੀ ਦੀਆਂ ਬਿਮਾਰੀਆਂ ਦੀ ਬਿਹਤਰ ਪਛਾਣ ਕਰ ਸਕਦੀ ਹੈ, ਹਾਲਾਂਕਿ ਇਸਨੇ ਕਿਸੇ ਵੀ ਚਿਤਰ ਨੂੰ ਜਿਸ ਨਾਲ ਲੰਬਾਈ-ਪੈਮਾਨਾ ਦਿਖਾਇਆ ਗਿਆ ਹੋਵੇ "ਕੈਂਸਰ" ਵਜੋਂ ਸ਼੍ਰੇਣੀਬੱਧ ਕੀਤਾ, ਕਿਉਂਕਿ ਮਾੜੀਆਂ ਬਿਮਾਰੀਆਂ ਦੀਆਂ ਤਸਵੀਰਾਂ ਨਾਲ ਆਮ ਤੌਰ 'ਤੇ ਐਸਾ ਕੀਤਾ ਜਾਂਦਾ ਹੈ। ਇਹ ਇਕ ਖਤਰਨਾਕ ਸਥਿਤੀ ਹੈ। ਇਕ ਹੋਰ ਖਤਰਨਾਕ ਉਦਾਹਰਣ 2015 ਵਿਚ ਰਿਚ ਕਾਰੂਆਨਾ ਦੁਆਰਾ ਲੱਭੀ ਗਈ ਸੀ: ਇਕ ਮਸ਼ੀਨ-ਲਰਨਿੰਗ ਪ੍ਰਣਾਲੀ ਜਿਸ ਨੇ ਮੌਤ ਦੇ ਖ਼ਤਰੇ ਦੀ ਸਹੀ ਭਵਿੱਖ-ਬਾਣੀ ਕੀਤੀ ਸੀ, ਨੇ 65 ਸਾਲ ਤੋਂ ਵੱਧ ਉਮਰ ਦੇ ਮਰੀਜ਼, ਦਮਾ ਅਤੇ ਸਾਹ ਲੈਣ ਵਿਚ ਮੁਸ਼ਕਲ ਨੂੰ "ਘੱਟ ਖ਼ਤਰਾ" ਵਜੋਂ ਸ਼੍ਰੇਣੀਬੱਧ ਕੀਤਾ ਸੀ। ਅੱਗੇ ਦੀ ਖੋਜ ਨੇ ਦਿਖਾਇਆ ਕਿ ਇਸ ਤਰ੍ਹਾਂ ਦੇ ਉੱਚ-ਖ਼ਤਰੇ ਵਾਲੇ ਮਾਮਲਿਆਂ ਵਿਚ, ਹਸਪਤਾਲ ਵਧੇਰੇ ਜਾਣਕਾਰੀ ਦੇਵੇਗਾ ਅਤੇ ਮਰੀਜ਼ ਦੀ ਜਾਨ ਬਚਾਏਗਾ ਅਤੇ ਪ੍ਰੋਗਰਾਮ ਦੁਆਰਾ ਮਾਪੇ ਗਏ ਖ਼ਤਰਿਆਂ ਨੂੰ ਘਟਾਏਗਾ। ਇਸ ਤਰ੍ਹਾਂ ਦੀਆਂ ਗਲਤੀਆਂ ਉਦੋਂ ਸਪੱਸ਼ਟ ਹੋ ਜਾਂਦੀਆਂ ਹਨ ਜਦੋਂ ਅਸੀਂ ਜਾਣਦੇ ਹਾਂ ਕਿ ਪ੍ਰੋਗਰਾਮ ਕਿਸੇ ਫੈਸਲੇ 'ਤੇ ਕਿਵੇਂ ਪਹੁੰਚਿਆ ਹੈ। ਬਿਨਾਂ ਕਿਸੇ ਸਪੱਸ਼ਟੀਕਰਨ ਦੇ, ਇਹਨਾਂ ਸਮੱਸਿਆਵਾਂ ਨੂੰ ਉਦੋਂ ਤਕ ਨਹੀਂ ਲੱਭਿਆ ਜਾ ਸਕਦਾ ਜਦੋਂ ਤਕ ਉਹ ਨੁਕਸਾਨ ਨਹੀਂ ਪਹੁੰਚਾਉਂਦੇ।

ਦੂਜਾ ਮੁੱਦਾ ਇਹ ਹੈ ਕਿ ਜਿਨ੍ਹਾਂ ਲੋਕਾਂ ਨੂੰ ਐਲਗੋਰਿਦਮ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀ) ਦੇ ਫੈਸਲੇ ਨਾਲ ਨੁਕਸਾਨ ਪਹੁੰਚਿਆ ਹੈ, ਉਨ੍ਹਾਂ ਨੂੰ ਸਪੱਸ਼ਟੀਕਰਨ ਦਾ ਅਧਿਕਾਰ ਹੈ। ਉਦਾਹਰਣ ਵਜੋਂ, ਡਾਕਟਰਾਂ ਨੂੰ ਆਪਣੇ ਕਿਸੇ ਵੀ ਫੈਸਲੇ ਦੇ ਪਿੱਛੇ ਦੇ ਤਰਕ (reasoning) ਨੂੰ ਸਪੱਸ਼ਟ ਕਰਨ ਦੀ ਅਤੇ ਪੂਰੀ ਤਰ੍ਹਾਂ ਸਮਝਾਉਣ ਦੀ ਲੋੜ ਹੁੰਦੀ ਹੈ। 2016 ਵਿਚ ਯੂਰਪੀਅਨ ਯੂਨੀਅਨ ਦੇ ਜਨਰਲ ਡਾਟਾ ਪ੍ਰੋਟੈਕਸ਼ਨ ਰੈਗੂਲੇਸ਼ਨ ਦੇ ਸ਼ੁਰੂਆਤੀ ਖਰੜੇ ਵਿਚ ਇਕ ਸਪੱਸ਼ਟ ਬਿਆਨ ਸ਼ਾਮਲ ਸੀ ਕਿ ਮਰੀਜ਼ਾਂ ਦਾ ਇਹ ਅਧਿਕਾਰ ਮੌਜੂਦ ਹੈ। ਉਦਯੋਗ ਮਾਹਿਰਾਂ ਨੇ ਨੋਟ ਕੀਤਾ ਕਿ ਇਹ ਇਕ ਅਣਸੁਲਝੀ ਸਮੱਸਿਆ ਹੈ ਜਿਸਦਾ ਕੋਈ ਹੱਲ ਨਜ਼ਰ ਨਹੀਂ ਆ ਰਿਹਾ ਹੈ। ਰੈਗੂਲੇਟਰਾਂ ਨੇ ਦਲੀਲ ਦਿੱਤੀ ਕਿ ਫਿਰ ਵੀ ਨੁਕਸਾਨ ਤਾਂ ਹੋ ਸਕਦਾ ਹੈ। ਜੇ ਸਮੱਸਿਆ ਦਾ ਕੋਈ ਹੱਲ ਨਹੀਂ ਹੈ, ਤਾਂ ਸਾਧਨਾਂ ਦੀ ਵਰਤੋਂ ਨਹੀਂ ਕੀਤੀ ਜਾਣੀ ਚਾਹੀਦੀ।

ਸਪੱਸ਼ਟਤਾ ਦੀ ਸਮੱਸਿਆ ਦੇ ਕਈ ਸੰਭਾਵੀ ਹੱਲ ਹਨ। 'ਮਲਟੀਟਾਸਕ (ਬਹੁਮੁਖੀ) ਲਰਨਿੰਗ' ਟੀਚੇ ਦੇ ਵਰਗੀਕਰਣ ਤੋਂ ਇਲਾਵਾ ਵੱਡੀ ਗਿਣਤੀ ਵਿਚ ਆਉਟਪੁੱਟ ਪ੍ਰਦਾਨ ਕਰਦੀ ਹੈ। ਇਹ ਤਰ੍ਹਾਂ ਦੀ ਹੋਰ ਆਉਟਪੁੱਟ ਡਿਵੈਲਪਰਾਂ ਨੂੰ ਇਹ ਪਤਾ ਲਗਾਉਣ ਵਿਚ ਮਦਦ ਕਰ ਸਕਦੀ ਹੈ ਕਿ ਨੈੱਟਵਰਕ ਨੇ ਕੀ ਸਿੱਖਿਆ ਹੈ। 'ਡੀਕਨਵੈਲਿਊਸ਼ਨ', 'ਡੀਪਡ੍ਰੀਮ' ਅਤੇ ਹੋਰ 'ਜਨਰੇਟਿਵ' ਤਰੀਕੇ ਡਿਵੈਲਪਰਾਂ ਨੂੰ ਇਹ ਦੇਖਣ ਦੀ ਆਗਿਆ ਦੇ ਸਕਦੇ ਹਨ ਕਿ ਡੂੰਘੇ ਨੈੱਟਵਰਕ ਦੀਆਂ ਵੱਖ-ਵੱਖ ਪਰਤਾਂ ਨੇ ਕੀ ਸਿੱਖਿਆ ਹੈ ਅਤੇ ਐਸੀ ਆਉਟਪੁੱਟ ਪੈਦਾ ਹੋ ਸਕਦੀ ਹੈ ਜੋ ਸੁਝਾਅ ਦੇਵੇ ਕਿ ਨੈੱਟਵਰਕ ਕੀ ਸਿੱਖ ਰਿਹਾ ਹੈ।

5. ਮਾੜੇ ਅਭਿਨੇਤਾ ਅਤੇ ਹਥਿਆਰਬੰਦ ਏਆਈ

ਘਾਤਕ ਖੁਦਮੁਖਤਿਆਰੀ ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਹਥਿਆਰ ਅਤੇ AI ਸੁਰੱਖਿਆ (defense) ਸਿਸਟਮ ਐਸੀਆਂ ਮਸ਼ੀਨਾਂ ਹਨ ਜੋ ਮਨੁੱਖੀ ਨਿਗਰਾਨੀ ਤੋਂ ਬਿਨਾਂ ਦੇਸ਼ਾਂ ਦੇ ਮਨੁੱਖੀ ਟੀਚਿਆਂ ਦਾ

ਪਤਾ ਲਗਾਉਂਦੀਆਂ ਹਨ ਅਤੇ ਉਨ੍ਹਾਂ ਨੂੰ ਲਾਗੂ ਕਰਦੀਆਂ ਹਨ। 2015 ਤਕ, ਪੰਜਾਹ ਤੋਂ ਵੱਧ ਦੇਸ਼ਾਂ ਨੂੰ ਜੰਗ ਦੇ ਮੈਦਾਨ ਦੇ ਰੋਬੋਟਾਂ ਦੀ ਖੋਜ ਕਰਨ ਦੀ ਰਿਪੋਰਟ ਕੀਤੀ ਗਈ ਸੀ। ਇਨ੍ਹਾਂ ਹਥਿਆਰਾਂ ਨੂੰ ਕਈ ਕਾਰਨਾਂ ਕਰਕੇ ਖਾਸ ਤੌਰ 'ਤੇ ਖ਼ਤਰਨਾਕ ਮੰਨਿਆ ਜਾਂਦਾ ਹੈ: ਜੇ ਐਸੇ ਹਥਿਆਰ ਕਿਸੇ ਨਿਰਦੋਸ਼ ਵਿਅਕਤੀ ਨੂੰ ਮਾਰਦੇ ਹਨ ਤਾਂ ਇਹ ਸਪੱਸ਼ਟ ਨਹੀਂ ਹੈ ਕਿ ਕਿਸ ਨੂੰ ਜਵਾਬਦੇਹ ਠਹਿਰਾਇਆ ਜਾਣਾ ਚਾਹੀਦਾ ਹੈ, ਇਹ ਵੀ ਸੰਭਾਵਨਾ ਨਹੀਂ ਹੈ ਕਿ ਉਹ ਭਰੋਸੇਯੋਗ ਤਰੀਕੇ ਨਾਲ ਨਿਸ਼ਾਨੇ ਚੁਣਨਗੇ, ਅਤੇ ਜੇ ਇਹ ਵੱਡੇ ਪੱਧਰ 'ਤੇ ਤਿਆਰ ਕੀਤੇ ਜਾਂਦੇ ਹਨ, ਤਾਂ ਉਹ ਸੰਭਾਵਿਤ ਤੌਰ 'ਤੇ ਸਮੂਹਿਕ ਵਿਨਾਸ਼ ਦੇ ਹਥਿਆਰ ਹਨ। 2014 ਵਿਚ, 30 ਦੇਸ਼ਾਂ (ਚੀਨ ਸਮੇਤ) ਨੇ ਕੁਝ ਰਵਾਇਤੀ ਹਥਿਆਰਾਂ 'ਤੇ ਸੰਯੁਕਤ ਰਾਸ਼ਟਰ ਦੀ ਕਨਵੈਨਸ਼ਨ ਦੇ ਤਹਿਤ ਖੁਦਮੁਖਤਿਆਰੀ ਹਥਿਆਰਾਂ 'ਤੇ ਪਾਬੰਦੀ ਦਾ ਸਮਰਥਨ ਕੀਤਾ, ਹਾਲਾਂਕਿ ਸੰਯੁਕਤ ਰਾਜ ਅਤੇ ਕਈ ਹੋਰ ਅਸਹਿਮਤ ਸਨ। ਇਸ ਮਾਮਲੇ ਵਿਚ ਅਸਲੀ ਰੂਪ ਵਿਚ ਸਹਿਮਤੀ ਦੀ ਸੰਭਾਵਨਾ ਨਹੀਂ। ਬਹੁਤ ਦੇਸ਼ ਐਸੇ ਹਥਿਆਰ ਬਣਾ ਰਹੇ ਹਨ ਤੇ ਬਣਾਉਂਦੇ ਰਹਿਣਗੇ। ਇਹ ਬਹੁਤ ਹੀ ਫ਼ਿਕਰ ਵਾਲੀ ਗਲ ਹੈ।

AI ਬਹੁਤ ਸਾਰੇ ਸਾਧਨ ਪ੍ਰਦਾਨ ਕਰਦਾ ਹੈ ਜੋ ਤਾਨਾਸ਼ਾਹੀ ਸਰਕਾਰਾਂ ਲਈ ਵਿਸ਼ੇਸ਼ ਤੌਰ 'ਤੇ ਲਾਭਦਾਇਕ ਹਨ: ਸਮਾਰਟ ਸਪਾਈਵੇਅਰ, ਚਿਹਰੇ ਦੀ ਪਛਾਣ ਅਤੇ ਆਵਾਜ਼ ਪਛਾਣ ਆਦਿ ਵਿਆਪਕ (ਬਹੁਤ ਤਰ੍ਹਾਂ ਦੀ) ਨਿਗਰਾਨੀ ਦੀ ਤਾਕਤ ਦਿੰਦੀ ਹੈ; ਅਜਿਹੀ ਨਿਗਰਾਨੀ ਮਸ਼ੀਨ-ਲਰਨਿੰਗ ਨੂੰ ਰਾਜ ਦੇ ਸੰਭਾਵਿਤ ਦੁਸ਼ਮਣਾਂ ਨੂੰ ਸ਼੍ਰੇਣੀਬੱਧ ਕਰਨ ਦੀ ਆਗਿਆ ਦਿੰਦੀ ਹੈ ਅਤੇ ਮਾੜੇ ਅਨਸਰਾਂ ਨੂੰ ਜਾਂ ਉਨ੍ਹਾਂ ਨੂੰ ਜੋ ਤੁਹਾਡੇ ਨਾਲ ਸਹਿਮਤ ਨਹੀਂ ਹਨ, ਨੂੰ ਕਾਬੂ ਕਰਨ ਵਿਚ ਸਹਾਇਤਾ ਦਿੰਦੀ ਹੈ। ਪਰ ਇਸ ਤਰ੍ਹਾਂ ਕਰਨ ਨਾਲ ਇਨਸਾਨਾਂ ਨੂੰ ਨਿੱਜੀ ਪਰਦੇਦਾਰੀ (ਪਰਵਿਓਚੇ) ਖਤਮ ਹੋ ਜਾਂਦੀ ਹੈ। AI ਦੀਆਂ ਸਿਫ਼ਾਰਿਸ਼ ਪ੍ਰਣਾਲੀਆਂ ਵੱਧ ਤੋਂ ਵੱਧ ਪ੍ਰਭਾਵ ਲਈ, ਪ੍ਰਚਾਰ ਲਈ ਅਤੇ ਗਲਤ ਜਾਣਕਾਰੀ ਨੂੰ ਸਹੀ ਢੰਗ ਨਾਲ ਨਿਸ਼ਾਨਾ ਬਣਾ ਸਕਦੀਆਂ ਹਨ (ਅਤੇ ਬਣਾ ਰਹੀਆਂ ਹਨ)। 'ਡੀਪਫੇਕਸ' ਅਤੇ 'ਜਨਰੇਟਿਵ' (ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕੰਪਿਊਟਰ ਸਮੱਗਰੀ ਬਣਾਉਣ ਵਾਲੀ) AI ਗਲਤ ਜਾਣਕਾਰੀ ਪੈਦਾ ਕਰਨ ਵਿਚ ਸਹਾਇਤਾ ਕਰਦੇ ਹਨ। ਉੱਨਤ AI ਤਾਨਾਸ਼ਾਹੀ ਕੇਂਦਰੀਕ੍ਰਿਤ ਫੈਸਲੇ ਲੈਣ ਨੂੰ ਮੁਕਾਬਲੇ ਵਿਚ ਵਧੇਰੇ ਤਾਕਤਵਰ (ਮੁਕਾਬਲੇਬਾਜ਼) ਬਣਾ ਸਕਦੀ ਹੈ, ਉਦਾਰਵਾਦੀ ਅਤੇ ਵਿਕੇਂਦਰੀਕ੍ਰਿਤ ਪ੍ਰਣਾਲੀਆਂ ਨਾਲੋਂ ਜਿਵੇਂ ਕਿ ਆਜ਼ਾਦ ਬਾਜ਼ਾਰਾਂ/ਵਪਾਰਾਂ/ਰਾਜਨੀਤੀ ਮੁਕਾਬਲਿਆਂ ਨਾਲ।

ਅੱਤਵਾਦੀ, ਅਪਰਾਧੀ ਅਤੇ ਦੁਸ਼ਟ ਦੇਸ਼ ਹਥਿਆਰਬੰਦ AI ਦੀ ਵਰਤੋਂ ਕਰ ਸਕਦੇ ਹਨ ਜਿਵੇਂ ਕਿ ਉੱਨਤ ਡਿਜੀਟਲ ਯੁੱਧ, ਖੁਦਮੁਖਤਿਆਰੀ ਹਥਿਆਰ ਅਤੇ ਉੱਨਤ ਰੋਬੋਟ ਹਥਿਆਰ। ਬੁਰੇ ਅਨਸਰ AI ਸਿਸਟਮ ਨੂੰ ਹੈਕ ਕਰਕੇ ਬਹੁਤ ਨੁਕਸਾਨ ਵੀ ਕਰ ਸਕਦੇ ਹਨ। ਐਸੀਆਂ ਹਰਕਤਾਂ ਦੇ ਨੁਕਸਾਨ ਦਾ ਅੰਦਾਜ਼ਾ ਲਗਾਉਣਾ ਔਖਾ ਹੈ।

ਮਸ਼ੀਨ-ਲਰਨਿੰਗ AI ਕੁਝ ਘੰਟਿਆਂ ਵਿਚ ਹਜ਼ਾਰਾਂ ਜ਼ਹਿਰੀਲੇ ਅਣੂਆਂ ਨੂੰ ਡਿਜ਼ਾਈਨ ਕਰਨ ਦੇ ਯੋਗ ਹੈ।

AI ਨਾਲ ਇੰਟਰਨੈੱਟ ਨੂੰ ਕਾਬੂ ਕਰਨ ਦੇ ਹਥਿਆਰ ਵੀ ਬਣਾਏ ਜਾ ਸਕਦੇ ਹਨ ਅਤੇ ਇਸ ਦੇ ਅਣਗਿਣਤ ਤੇ ਘਾਤਕ ਨਤੀਜੇ ਹੋ ਸਕਦੇ ਹਨ।

6 ਤਕਨੀਕੀ ਬੇਰੁਜ਼ਗਾਰੀ

AI ਦੀ ਮਦਦ ਨਾਲ ਬਹੁਤ ਸਾਰੇ ਕੰਪਿਊਟਰਾਂ ਦੁਆਰਾ ਕੀਤੇ ਗਏ ਕੰਮ ਤੇ ਕਈ ਹੋਰ ਕੰਮ ਬੜੀ ਆਸਾਨੀ ਨਾਲ ਵੱਢੀ ਮਿਕਦਾਰ ਵਿਚ ਥੋੜ੍ਹੇ ਇਨਸਾਨਾਂ ਨਾਲ ਕੀਤੇ ਜਾ ਸਕਦੇ ਹਨ। ਇਸ ਦੇ ਕਈ ਗੰਭੀਰ ਅਸਰ ਹੋ ਸਕਦੇ ਹਨ ਤੇ ਬੇਰੁਜ਼ਗਾਰੀ ਵਧ ਸਕਦੀ ਹੈ। ਦੂਜੇ ਪਾਸੇ ਜਿਨ੍ਹਾਂ ਕੰਪਨੀਆਂ ਕੋਲ AI ਦੀ ਤਾਕਤ ਹੈ ਉਨ੍ਹਾਂ ਦੀ ਕਮਾਈ ਬਹੁਤ ਵਧ ਸਕਦੀ ਹੈ।

ਹੁਣ ਤਕ ਤਾਂ ਤਕਨਾਲੋਜੀ ਨੇ ਕੁੱਲ ਰੁਜ਼ਗਾਰ ਨੂੰ ਘਟਾਉਣ ਦੀ ਬਜਾਏ ਵਧਾਉਣ ਦਾ ਰੁਝਾਨ ਰੱਖਿਆ ਹੈ, ਪਰ ਅਰਥ-ਸ਼ਾਸਤਰੀ ਮੰਨਦੇ ਹਨ ਕਿ "ਅਸੀਂ AI ਦੇ ਨਾਲ ਅਣਜਾਣ ਖੇਤਰ ਵਿਚ ਹਾਂ"। ਅਰਥ-ਸ਼ਾਸਤਰੀਆਂ ਦੇ ਇਕ ਸਰਵੇਖਣ ਨੇ ਇਸ ਬਾਰੇ ਅਸਹਿਮਤੀ ਦਿਖਾਈ ਕਿ ਕੀ ਰੋਬੋਟਾਂ ਅਤੇ AI ਦੀ ਵਧਦੀ ਵਰਤੋਂ ਲੰਬੇ ਸਮੇਂ ਦੀ ਬੇਰੁਜ਼ਗਾਰੀ ਵਿਚ ਕਾਫ਼ੀ ਵਾਧਾ ਕਰੇਗੀ, ਪਰ ਉਹ ਆਮ ਤੌਰ 'ਤੇ ਸਹਿਮਤ ਹੁੰਦੇ ਹਨ ਕਿ ਜੇ ਉਤਪਾਦਕਤਾ ਦੇ ਲਾਭਾਂ ਨੂੰ ਦੁਬਾਰਾ ਵੰਡਿਆ ਜਾਂਦਾ ਹੈ ਤਾਂ ਇਹ ਸ਼ੁੱਧ ਲਾਭ ਹੋ ਸਕਦਾ ਹੈ।

ਆਟੋਮੇਸ਼ਨ ਦੀਆਂ ਪਿਛਲੀਆਂ ਲਹਿਰਾਂ ਦੇ ਦੌਰਾਨ, ਬਹੁਤ ਸਾਰੀਆਂ ਮੱਧ-ਵਰਗੀ ਨੌਕਰੀਆਂ ਨੂੰ ਆਰਟੀਫ਼ੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਦੁਆਰਾ ਖਤਮ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ; "ਦੀ ਇਕੋਨਾਮਿਸਟ" ਨੇ 2015 ਵਿਚ

ਕਿਹਾ ਸੀ ਕਿ ਇਹ ਚਿੰਤਾ ਹੈ ਕਿ AI ਵਾਈਟ-ਕਾਲਰ ਨੌਕਰੀਆਂ ਲਈ ਉਹੀ ਕਰ ਸਕਦੀ ਹੈ ਜੋ ਉਦਯੋਗਿਕ ਕ੍ਰਾਂਤੀ ਦੌਰਾਨ ਭਾਫ਼ ਸ਼ਕਤੀ ਨੇ ਬਲੂ-ਕਾਲਰ ਨੌਕਰੀਆਂ ਨਾਲ ਕੀਤਾ ਸੀ। ਇਹ ਇਕ ਗੰਭੀਰਤਾ ਵਾਲਾ ਵਿਸ਼ਾ ਹੈ। ਬਹੁਤ ਜ਼ਿਆਦਾ ਖ਼ਤਰੇ ਵਾਲੀਆਂ ਨੌਕਰੀਆਂ ਹਨ ਵਕੀਲ-ਸਹਾਇਤਾ ਕਰਨ ਵਾਲਿਆਂ ਤੋਂ ਲੈ ਕੇ ਫਾਸਟ-ਫੂਡ (ਛੇਤੀ-ਭੋਜਨ) ਰਸੋਈਏ ਤਕ। ਸਿਹਤ-ਸੰਭਾਲ ਵਿਧੀਆਂ ਵਿਚ ਵੀ ਕਈ ਤਬਦੀਲੀਆਂ ਆਉਣਗੀਆਂ ਕਿਉਂਕਿ ਇਸ ਵਿਚ ਕਈ ਕੰਮ AI ਨੂੰ ਸੌਂਪ ਦਿੱਤੇ ਜਾਣਗੇ।

ਅਪ੍ਰੈਲ 2023 ਵਿਚ, ਇਹ ਦੱਸਿਆ ਗਿਆ ਸੀ ਕਿ ਚੀਨੀ ਵੀਡੀਓ ਖੇਲ ਚਿੱਤਰਕਾਰਾਂ ਦੀਆਂ 70% ਨੌਕਰੀਆਂ ਜਨਰੇਟਿਵ (ਕਈ ਤਰ੍ਹਾਂ ਦੀ ਕੰਪਿਊਟਰ ਸਮੱਗਰੀ ਬਣਾਉਣ ਵਾਲੀ) ਆਰਟੀਫੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਦੁਆਰਾ ਖਤਮ ਕਰ ਦਿੱਤੀਆਂ ਗਈਆਂ ਸਨ।

7. ਮਨੁੱਖੀ ਹੋਂਦ ਦਾ ਖ਼ਤਰਾ

ਇਹ ਦਲੀਲ ਦਿੱਤੀ ਗਈ ਹੈ ਕਿ AI ਇੰਨੀ ਸ਼ਕਤੀਸ਼ਾਲੀ ਹੋ ਜਾਵੇਗੀ ਕਿ ਮਨੁੱਖਤਾ ਇਸ 'ਤੇ ਨਿਯੰਤਰਣ/ਕਾਬੂ ਗੁਆ ਸਕਦੀ ਹੈ। ਜਿਵੇਂ ਕਿ ਭੌਤਿਕ ਵਿਗਿਆਨੀ ਸਟੀਫਨ ਹਾਕਿੰਗ ਕਹਿੰਦੇ ਹਨ, ਇਹ "ਮਨੁੱਖ ਜਾਤੀ ਦੇ ਅੰਤ ਦਾ ਸੰਕੇਤ ਦੇ ਸਕਦਾ ਹੈ"। ਇਹ ਦ੍ਰਿਸ਼ ਸਾਇੰਸ ਫਿਕਸ਼ਨ/ਮਿਥਿਹਾਸ ਵਿਚ ਆਮ ਰਿਹਾ ਹੈ, ਜਦੋਂ ਇਕ ਕੰਪਿਊਟਰ ਜਾਂ ਰੋਬੋਟ ਅਚਾਨਕ ਮਨੁੱਖ ਵਰਗੀ "ਸਵੈ-ਜਾਗਰੂਕਤਾ" (ਜਾਂ "ਸੰਵੇਦਨਸ਼ੀਲਤਾ" ਜਾਂ "ਚੇਤਨਾ") ਵਿਕਸਤ ਕਰਦਾ ਹੈ ਅਤੇ ਇਕ ਮਾੜਾ ਕਿਰਦਾਰ ਬਣ ਜਾਂਦਾ ਹੈ। ਇਹ ਸਾਇੰਸ ਦਰਸਾਉਣ ਵਾਲਾ ਦ੍ਰਿਸ਼ ਕਈ ਤਰੀਕਿਆਂ ਨਾਲ ਗੁੰਮਰਾਹ ਕਰਨ ਵਾਲਾ ਹੈ। ਪਰ ਉਸ ਦੇ ਉਲਟ ਕਈ ਅਸਲੀ ਤੇ ਗੰਭੀਰ ਖ਼ਤਰੇ ਨਜ਼ਰ ਆਉਂਦੇ ਹਨ।

ਸਭ ਤੋਂ ਪਹਿਲਾਂ, AI ਨੂੰ ਹੋਂਦ ਦੀ ਤਬਾਹੀ ਵਜੋਂ ਮਨੁੱਖ ਵਰਗੀ "ਸੰਵੇਦਨਾਸ਼ੀਲਤਾ ਜਾਂ ਚੇਤਨਾ" ਦੀ ਜ਼ਰੂਰਤ ਨਹੀਂ ਹੈ। ਆਧੁਨਿਕ AI ਪ੍ਰੋਗਰਾਮਾਂ ਨੂੰ ਵਿਸ਼ੇਸ਼ ਟੀਚੇ ਦਿੱਤੇ ਜਾਂਦੇ ਹਨ ਅਤੇ ਉਨ੍ਹਾਂ ਨੂੰ ਪ੍ਰਾਪਤ ਕਰਨ ਲਈ ਸਿੱਖਣ ਅਤੇ ਬੁੱਧੀ ਦੀ ਵਰਤੋਂ ਕੀਤੀ ਜਾਂਦੀ ਹੈ। ਦਾਰਸ਼ਨਿਕ ਨਿਕ ਬੋਸਟਰੋਮ ਨੇ ਦਲੀਲ ਦਿੱਤੀ ਕਿ ਜੇ ਕੋਈ ਕਾਫ਼ੀ ਸ਼ਕਤੀਸ਼ਾਲੀ AI ਨੂੰ ਲਗਭਗ ਕੋਈ ਟੀਚਾ ਦਿੰਦਾ ਹੈ, ਤਾਂ ਇਹ ਇਸ ਨੂੰ ਪ੍ਰਾਪਤ ਕਰਨ ਲਈ ਮਨੁੱਖਤਾ ਨੂੰ ਤਬਾਹ ਕਰਨ ਦੀ ਚੋਣ ਕਰ ਸਕਦੀ ਹੈ। ਸਟੂਅਰਟ ਰਸਲ ਘਰੇਲੂ ਰੋਬੋਟ ਦੀ ਉਦਾਹਰਣ ਦਿੰਦਾ ਹੈ ਜੋ ਆਪਣੇ ਮਾਲਕ ਨੂੰ ਮਾਰਨ ਦਾ ਤਰੀਕਾ ਲੱਭਣ ਦੀ ਕੋਸ਼ਿਸ਼ ਕਰਦਾ ਹੈ ਤਾਂ ਜੋ ਉਹ ਮਾਲਿਕ ਦੇ ਹੁਕਮਾਂ ਤੋਂ ਆਜ਼ਾਦ ਹੋ ਸਕੇ। ਐਸਾ ਹੋਣਾ ਸ਼ਾਇਦ ਸੰਭਵ ਨਹੀਂ, ਪਰ ਜੇ ਕੋਈ ਰੋਬੋਟ ਨੂੰ ਮਾੜੇ ਕੰਮ ਲਈ ਗੁਲਾਮ ਬਣਾ ਲਏ ਤਾਂ ਪਤਾ ਨਹੀਂ ਕੀ ਕੁਝ ਵਿਨਾਸ਼ਕਾਰੀ ਨੁਕਸਾਨ ਕੀਤਾ ਜਾ ਸਕਦਾ।

ਦੂਜਾ, ਯੁਵਲ ਨੋਹ ਹਰਾਰੀ ਦਲੀਲ ਦਿੰਦਾ ਹੈ ਕਿ AI ਨੂੰ ਹੋਂਦ ਦਾ ਖ਼ਤਰਾ ਪੈਦਾ ਕਰਨ ਲਈ AI ਰੋਬੋਟ ਲਈ ਚਲਦੇ-ਫਿਰਦੇ ਸਰੀਰ ਦੀ ਜ਼ਰੂਰਤ ਨਹੀਂ ਹੈ। ਸਭਿਅਤਾ ਦੇ ਜ਼ਰੂਰੀ ਅੰਗ ਭੌਤਿਕ ਨਹੀਂ ਹਨ। ਵਿਚਾਰਧਾਰਾ, ਕਾਨੂੰਨ, ਸਰਕਾਰ, ਪੈਸਾ ਅਤੇ ਆਰਥਿਕਤਾ ਵਰਗੀਆਂ ਚੀਜ਼ਾਂ ਭਾਸ਼ਾ ਤੋਂ ਬਣੀਆਂ ਹੁੰਦੀਆਂ ਹਨ; ਉਹ ਇਸ ਲਈ ਮੌਜੂਦ ਹਨ ਕਿਉਂਕਿ ਅਜਿਹੀਆਂ ਕਹਾਣੀਆਂ ਹਨ ਜਿਨ੍ਹਾਂ 'ਤੇ ਅਰਬਾਂ ਲੋਕ ਵਿਸ਼ਵਾਸ ਕਰਦੇ ਹਨ। ਗਲਤ ਜਾਣਕਾਰੀ ਦਾ ਮੌਜੂਦਾ ਪ੍ਰਸਾਰ ਸੁਝਾਅ ਦਿੰਦਾ ਹੈ ਕਿ ਇਕ AI ਲੋਕਾਂ ਨੂੰ ਕਿਸੇ ਵੀ ਚੀਜ਼ 'ਤੇ ਵਿਸ਼ਵਾਸ ਕਰਨ ਲਈ ਯਕੀਨ ਦਿਵਾਉਣ ਲਈ ਭਾਸ਼ਾ ਦੀ ਵਰਤੋਂ ਕਰ ਸਕਦਾ ਹੈ, ਇਥੋਂ ਤਕ ਕਿ ਵਿਨਾਸ਼ਕਾਰੀ ਕਾਰਵਾਈਆਂ ਕਰਨ ਲਈ ਵੀ। ਸ਼ਾਇਦ ਸਭ ਤੋਂ ਖ਼ਤਰੇ ਵਾਲੀ ਸਥਿਤੀ ਇਹ ਹੈ ਕਿ ਲੋਕਾਂ ਨੂੰ AI ਰਾਹੀਂ ਮਾੜੇ ਕੰਮਾਂ ਲਈ ਪ੍ਰੇਰਿਤ ਕੀਤਾ ਜਾ ਸਕਦਾ ਹੈ।

ਮਾਹਿਰਾਂ ਅਤੇ ਉਦਯੋਗ ਦੇ ਅੰਦਰੂਨੀ ਲੋਕਾਂ ਵਿਚਾਲੇ ਵਿਚਾਰ ਮਿਸ਼ਰਿਤ ਹਨ, ਜਿਸ ਵਿਚ ਬਹੁਤ ਸਾਰੇ ਹਿੱਸੇ ਚਿੰਤਤ ਹਨ ਅਤੇ ਕੁਝ ਹੋ ਸਕਣ ਵਾਲੇ ਸੁਪਰਇੰਟੈਲੀਜੈਂਟ (ਬੇਅੰਤ ਬੁੱਧੀ ਵਾਲੀ) AI ਦੇ ਖ਼ਤਰਿਆਂ ਤੋਂ ਬੇਪਰਵਾਹ ਹਨ। ਸਟੀਫਨ ਹਾਕਿੰਗ, ਬਿਲ ਗੇਟਸ, ਐਲਨ ਮਸਕ ਵਰਗੀਆਂ ਸ਼ਖਸੀਅਤਾਂ ਨੇ AI ਤੋਂ ਹੋਂਦ ਦੇ ਖ਼ਤਰੇ ਬਾਰੇ ਚਿੰਤਾ ਜ਼ਾਹਰ ਕੀਤੀ ਹੈ। 2010 ਦੇ ਦਹਾਕੇ ਦੇ ਸ਼ੁਰੂ ਵਿਚ, ਮਾਹਰਾਂ ਨੇ ਦਲੀਲ ਦਿੱਤੀ ਕਿ ਭਵਿੱਖ ਵਿਚ ਦੇ ਖ਼ਤਰੇ ਬਹੁਤ ਦੂਰ ਹਨ ਕਿ ਖੋਜ ਦੀ ਇਹ ਗਰੰਟੀ ਨਹੀਂ ਦਿੱਤੀ ਜਾ ਸਕਦੀ ਕਿ ਮਨੁੱਖ ਇਕ ਸੁਪਰਇੰਟੈਲੀਜੈਂਟ ਮਸ਼ੀਨ ਦੇ ਨਜ਼ਰੀਏ ਤੋਂ ਜਾਣੂ ਹੋਣਗੇ। ਹਾਲਾਂਕਿ, 2016 ਤੋਂ ਬਾਅਦ, ਮੌਜੂਦਾ ਅਤੇ ਭਵਿੱਖ ਦੇ ਖ਼ਤਰਿਆਂ ਅਤੇ ਸੰਭਾਵਿਤ ਹੱਲਾਂ ਦਾ ਅਧਿਐਨ ਖੋਜ ਦਾ ਇਕ ਗੰਭੀਰ ਖੇਤਰ ਬਣ ਗਿਆ। 2023 ਵਿਚ, ਜੈਫਰੀ ਹਿੰਟਨ, ਯੋਸ਼ੁਆ ਬੇਂਜੀਓ, ਡੇਮਿਸ ਹਸਾਬਿਸ ਅਤੇ ਸੈਮ ਆਲਟਮੈਨ ਸਮੇਤ AI ਦੇ ਮੁੱਢਲੇ ਆਗੂਆਂ ਨੇ ਸੰਯੁਕਤ ਬਿਆਨ ਜਾਰੀ ਕੀਤਾ ਕਿ "AI ਦੇ ਗੁਪਤ ਰਹਿਣ ਵਾਲੇ ਖ਼ਤਰਿਆਂ ਨੂੰ ਘਟਾਉਣਾ, ਮਹਾਂ-ਮਾਰੀ ਅਤੇ

ਪ੍ਰਮਾਣੂ ਯੁੱਧ ਵਰਗੇ ਹੋਰ ਸਮਾਜਕ ਪੱਧਰ ਦੇ ਖ਼ਤਰਿਆਂ ਦੇ ਨਾਲ ਇਕ ਵਿਸ਼ਵ-ਵਿਆਪੀ ਤਰਜੀਹ ਹੋਣੀ ਚਾਹੀਦੀ ਹੈ।"

ਨਿਚੋੜ

AI ਇਕ ਬਹੁਤ ਸ਼ਕਤੀਸ਼ਾਲੀ ਸੰਦ ਹੈ ਜਿਸ ਦੇ ਪੂਰੇ ਅਸਰ ਸਾਨੂੰ ਪਤਾ ਵੀ ਨਹੀਂ ਹਨ। ਹੁਣ ਤਕ ਦੇ ਵੀ ਇਸ ਦੇ ਕਈ ਸ਼ਕਤੀਸ਼ਾਲੀ ਪ੍ਰਭਾਵ ਨਜ਼ਰ ਆਉਂਦੇ ਹਨ। ਇਹ ਤਾਂ ਸਾਫ਼ ਹੈ ਕਿ ਇਸ ਦੇ ਬਹੁਤ ਫਾਇਦੇ ਹਨ ਅਤੇ ਕਈ ਨੁਕਸਾਨ ਵੀ ਹਨ। ਇੰਟਰਨੈੱਟ ਰਾਹੀਂ ਉਪਲੱਭਧ ਸਮਾਜਕ/ਸੋਸ਼ਲ-ਮੀਡਿਆ ਨੇ ਪਹਿਲਾਂ ਹੀ ਬਹੁਤ ਸਮਾਜਕ ਬਦਲਾਵ ਕੀਤੇ ਹਨ ਜਿਨ੍ਹਾਂ ਵਿਚੋਂ ਇਕ ਝੂਠ ਤੇ ਪ੍ਰਾਪੇਗੰਡਾ ਫੈਲਾਉਣ ਦਾ ਸੌਖਾ ਹੋਣਾ ਸੀ, ਪਰ AI ਦੇ ਐਲਗੋਰਿਦਮਾ (ਨਿਯਮੋਂ ਦੀ ਪਰਨਾਲੀਆਂ) ਨਾਲ ਹੋਰ ਵੀ ਸੌਖਾ ਹੋ ਗਿਆ ਹੈ। ਸੋਸ਼ਲ-ਮੀਡਿਆ ਨੇ ਜੋ ਕਈ ਪੱਧਰਾਂ ਦੀਆਂ ਸਮਾਜਕ ਦਰਾੜਾਂ ਪਾਈਆਂ ਸੀ ਉਹ ਹੋਰ ਡੂੰਘੀਆਂ ਹੋ ਰਹੀਆਂ ਹਨ। ਸਭ ਤੋਂ ਘਾਟੇ ਵਾਲੀ ਸਥਿਤੀ ਇਹ ਬਣ ਸਕਦੀ ਹੈ ਕਿ ਲੋਕ ਕਿਸੇ ਖ਼ਬਰ ਜਾਂ ਘਟਨਾ ਜਾਂ ਹਕੀਕਤ ਦਾ ਵਿਸ਼ਵਾਸ ਕਰਨਾ ਛੱਡ ਦੇਣ ਇਸ ਦੀ ਸੰਭਾਵਨਾ ਹੈ।

ਤਾਨਾਸ਼ਾਹੀ ਜਾਂ ਰਿਸ਼ਵਤਖੋਰ ਸਰਕਾਰਾਂ AI ਨੂੰ ਵਰਤ ਕੇ ਬਹੁਤ ਸਮਾਜਕ ਨੁਕਸਾਨ ਕਰ ਸਕਦੀਆਂ ਹਨ। ਇਸ ਤਰ੍ਹਾਂ ਕਰਨ ਨਾਲ ਸਮਾਜਕ ਅਣਬਣ ਹੋ ਸਕਦੀ ਹੈ, ਲੜਾਈ ਝਗੜੇ ਵੀ ਹੋ ਸਕਦੇ ਹਨ ਅਤੇ ਕ੍ਰਾਂਤੀਕਾਰੀ ਅੰਦੋਲਨ ਵੀ।

ਹੁਣ ਤਕ ਦੀ AI ਰਾਹੀਂ ਸਮਾਜਕ ਪੱਖਪਾਤ, ਨਸਲ, ਲਿੰਗ, ਲਿੰਗਕਤਾ, ਅਤੇ ਨਸਲੀ ਵਿਤਕਰੇ ਹੋ ਸਕਦੇ ਹਨ।

AI ਦੀ ਮਦਦ ਨਾਲ ਬਹੁਤ ਸਾਰੇ ਕੰਪਿਊਟਰਾਂ ਦੁਆਰਾ ਕੀਤੇ ਗਏ ਕੰਮ ਤੇ ਕਈ ਹੋਰ ਕੰਮ ਬੜੀ ਆਸਾਨੀ ਨਾਲ ਵੱਡੀ ਮਿਕਦਾਰ ਵਿਚ ਥੋੜ੍ਹੇ ਇਨਸਾਨਾਂ ਨਾਲ ਕੀਤੇ ਜਾ ਸਕਦੇ ਹਨ ਅਤੇ ਇਸ ਤਰ੍ਹਾਂ ਬੇਰੁਜ਼ਗਾਰੀ ਵਧ ਸਕਦੀ ਹੈ। ਇਸ ਨਾਲ ਅਮੀਰ ਹੋਰ ਅਮੀਰ ਹੋ ਸਕਦੇ ਹਨ ਤੇ ਗਰੀਬ ਹੋਰ ਗਰੀਬ।

ਕਈ ਨਤੀਜਿਆਂ ਲਈ ਇਹ ਜਾਨਣਾ ਜ਼ਰੂਰੀ ਹੈ ਕਿ ਉਹ ਕੋਈ ਫੈਸਲਾ ਕਿਵੇਂ ਕੀਤਾ ਗਿਆ, ਪਰ ਇਹ ਜਾਨਣਾ ਮੁਸ਼ਕਲ ਹੋ ਗਿਆ ਹੈ।

ਜਿਨ੍ਹਾਂ ਕੋਲ AI ਦੀ ਸਹੂਲਤ ਹੈ ਜਾਂ ਹੋਵੇਗੀ ਉਹ ਦੂਜਿਆਂ ਨਾਲੋਂ ਕਈ ਖੇਤਰਾਂ ਵਿਚ ਬਹੁਤ ਸ਼ਕਤੀਸ਼ਾਲੀ ਹੋ ਜਾਣਗੇ। ਘਾਤਕ ਖੁਦਮੁਖਤਿਆਰੀ ਹਥਿਆਰ ਅਤੇ ਆਰਟੀਫੀਸ਼ੀਅਲ ਇੰਟੈਲੀਜੈਂਸ ਹਥਿਆਰਾਂ ਦੀ ਦੌੜ ਮਨੁੱਖਤਾ ਲਈ ਬਹੁਤ ਵੱਡੇ ਖ਼ਤਰੇ ਹਨ।

ਹਵਾਲੇ :

1. https://en.wikipedia.org/wiki/Artificial-intelligence (Wikipedia)
2. https://youtu.be/CC2W3KhaBsM?si=XwQ-vn7k6luIKzly
3. https://youtu.be/sitHS6UDMJc?si=2Qidy60Qo0xRbEUC
4. https://youtu.be/fLRc62-F1o0?si=2v0DmgjHpYYtmnaQ
5. https://youtu.be/rGgGOccMEiY?si=g4O3fUvKyVg577kt
6. https://youtu.be/qrvK-KuIeJk?si=NphsFBakpvU36ZAj
7. https://youtu.be/FAbsoxQtUwM?si=JmpJCdro0PXW6iuk
8. https://youtu.be/S4bBjoYeUAg?si=Y-Vv2S0mqvKDg8Gw
9. https://youtu.be/xUNx-PxNHrY?si=wXPEU7amRQsDFKOI
10. https://www.simplilearn.com/advantages-and-disadvantages-of-artificial-intelligence-article#disadvantages-of-artificial-intelligence