Usare un linguaggio di modellizzazione del linguaggio (LLM) come ChatGPT in modalità zero-shot significa sfruttare il modello per compiti senza fornire esempi specifici durante l'addestramento. In questa modalità, il modello cerca di generare una risposta basandosi solo sulla descrizione del compito fornita dall'utente.

Le performance di accuracy che si avvicinano a quelle di un algoritmo supervisionato potrebbero essere dovute a diversi fattori:

1. **Capacità di generalizzazione**: I LLM come ChatGPT sono addestrati su un vasto corpus di testo che copre una vasta gamma di argomenti e stili linguistici. Questo addestramento ampio consente loro di avere una buona comprensione della struttura del linguaggio e di poter fare inferenze ragionevoli anche su compiti non visti in fase di addestramento.
2. **Trasferimento di conoscenza**: Anche se non sono stati addestrati specificamente per un compito, i LLM come ChatGPT possono sfruttare la conoscenza appresa durante l'addestramento per affrontare nuovi compiti. Ciò è possibile grazie alla capacità del modello di catturare e utilizzare modelli di linguaggio e relazioni semantiche presenti nei dati di addestramento.
3. **Architettura complessa**: Gli LLM come ChatGPT sono modelli neurali profondi con un'architettura complessa che consente loro di apprendere rappresentazioni linguistiche di alto livello. Questa complessità permette loro di catturare relazioni complesse nel linguaggio, che possono essere utili per affrontare una varietà di compiti.
4. **Addestramento su dati strutturati e non strutturati**: Durante l'addestramento, i modelli come ChatGPT hanno accesso a una combinazione di dati strutturati e non strutturati. Questo consente loro di apprendere da una vasta gamma di fonti, inclusi testi annotati e non annotati, il che può contribuire alla loro capacità di svolgere compiti senza supervisione.

In sintesi, le performance simili a quelle di un algoritmo supervisionato in modalità zero-shot possono essere attribuite alla combinazione di una buona capacità di generalizzazione, trasferimento di conoscenza, architettura complessa e addestramento su dati diversificati. Tuttavia, è importante notare che le prestazioni possono variare a seconda del compito specifico e delle condizioni di utilizzo.