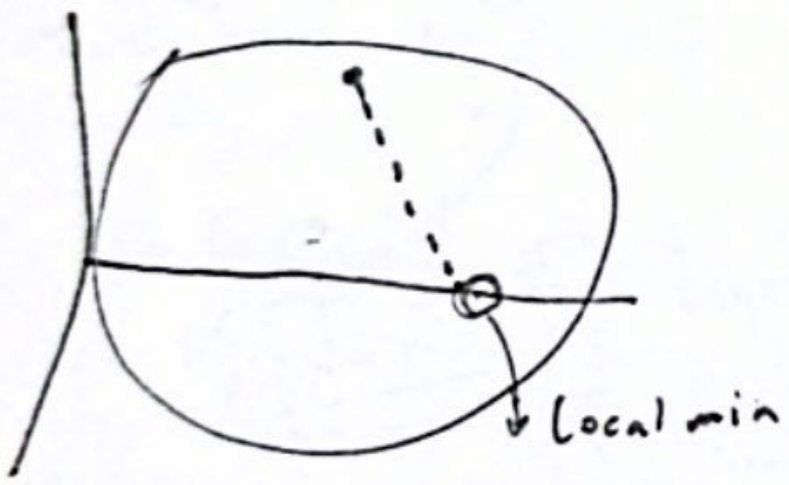
NN cont

Stochastic Gradient Decent

- With normal GD we need to eval every training img and add up all ith influence if we find the cost of each training image then to take it then to one decent step with GD and repete this again and again until some local min is reached.

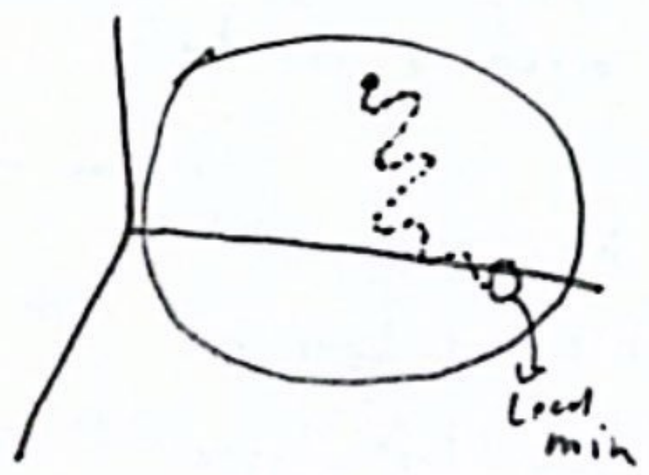
But this is slow

- So instead and more commonly we randomly shuffel training data the divide it into batches the we do a single step if eval one batch and apply Gradient Decent / Backprop once for this batch and we repete this process for all batches, this is much faster



Gradient Decent

Carefully calc step in steepest Direction



Stochastic Gradient Decent

faster less calculated steps but converging on local min at end

---

NLP (Natural language processing) is the intersection of CS, AI and linguistics

goal: to enable computers to understand interprit and generate human language

Ex: • text classification (NB)
• Machine translation
• sentiament analysis
• text summarization

LLMs are large scale implimentations of NLP
LLMs unify all NLP tasks under one thing LLM