মেশিন লার্নিং নোট (Machine Learning Note)

টপিক: লিনিয়ার রিগ্রেশন থেকে লজিস্টিক রিগ্রেশনে রূপান্তর (*Transition from Linear to Logistic Regression*)

১. লিনিয়ার রিগ্রেশন (Linear Regression)

লেকচারের প্রথম অংশে মাল্টিপল লিনিয়ার রিগ্রেশনের (Multiple Linear Regression) মূল মডেলটি আলোচনা করা হয়েছে। এখানে বেশ কয়েকটি ইনপুট ফিচার এবং তাদের সংশ্লিষ্ট ওয়েট (Weight) গুণ করে একটি বায়াস (Bias) যোগ করা হয়। গাণিতিক রূপটি নিম্নরূপ:

$$f_{w,b}(x) = w \cdot x + b = w_1x_1 + w_2x_2 + w_3x_3 + \dots + w_nx_n + b$$

- **আউটপুট পরিসর:** এই সমীকরণের আউটপুট যেকোনো বাস্তব সংখ্যা হতে পারে, যা পজিটিভ বা নেগেটিভ ইনফিনিটি (-∞ থেকে +∞) পর্যন্ত বিস্তৃত।
- **সীমাবদ্ধতা:** আমরা যখন কোনো ক্লাসিফিকেশন বা শ্রেণীবিভাগের কাজ করতে চাই (যেমন: একটি ইমেইল স্প্যাম (1) নাকি স্প্যাম নয় (0) তা প্রেডিক্ট করা), তখন এই সাধারণ লিনিয়ার আউটপুট সরাসরি ব্যবহার করা যায় না। কারণ এর মান ০ বা ১ এর চেয়ে অনেক বড় কিংবা ছোট হতে পারে, যা সরাসরি সম্ভাব্যতা (Probability) নির্দেশ করে না।

২. লজিস্টিক রিগ্রেশনে রূপান্তর (Transition to Logistic Regression)

লিনিয়ার রিগ্রেশনের আউটপুটকে ক্লাসিফিকেশনের উপযোগী করার জন্য **লজিস্টিক রিগ্রেশন** মডেলের অবতারণা করা হয়। এখানে লিনিয়ার রিগ্রেশনের প্রাপ্ত আউটপুটকে (যেটিকে আমরা z ধরি) একটি বিশেষ অ্যাক্টিভেশন ফাংশনের ভেতর দিয়ে পাস করানো হয়। এই ফাংশনটিকে বলা হয় **সিগময়েড ফাংশন (Sigmoid Function)**, যা গ্রীক অক্ষর সিগমা (σ) দ্বারা প্রকাশ করা হয়।

$$f_{w,b}(x) = \sigma(z)$$

যেখানে, z হলো আমাদের পূর্ববর্তী লিনিয়ার রিগ্রেশনের সমীকরণটি:

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b$$

৩. সিগময়েড ফাংশন এবং এর গাণিতিক রূপ

বোর্ডের পরবর্তী চিত্রগুলোতে সিগময়েড ফাংশন বা $\sigma(z)$-এর বীজগণিতীয় রূপটি ভেঙে দেখানো হয়েছে। সিগময়েড ফাংশনের মূল সূত্রটি হলো:

$$\sigma(z) = 1 / (1 + e^{-z})$$

এখন, z-এর স্থানে লিনিয়ার কম্বিনেশনের পুরো মানটি প্রতিস্থাপন করলে আমরা লজিস্টিক রিগ্রেশনের সম্পূর্ণ প্রেডিকশন বা হাইপোথিসিস সমীকরণটি পাই:

$$f_{w,b}(x) = 1 / (1 + e^{-(w_1x_1 + w_2x_2 + \dots + w_nx_n + b)})$$

কেন আমরা সিগময়েড ফাংশন ব্যবহার করি?

- **সম্ভাব্যতা ম্যাপিং (Probability Mapping):** z-এর মান ধনাত্মক বা ঋণাত্মক যতই বড় হোক না কেন, অয়লারের ধ্রুবক (e^{-z}) সম্বলিত এই ফাংশনটি চূড়ান্ত আউটপুটকে সবসময় কঠোরভাবে **০ থেকে ১ (0 to 1)** এর মধ্যে সীমাবদ্ধ রাখে। এই মানটিকে সরাসরি কোনো ক্লাসের অন্তর্ভুক্ত হওয়ার সম্ভাব্যতা (Probability) হিসেবে গণ্য করা হয়।
- **ডিসিশন বাউন্ডারি (Decision Boundary):** সাধারণত আউটপুট যদি ≥ **০.৫** হয়, তবে মডেলটি প্রেডিকশন হিসেবে ক্লাস 1 নির্ধারণ করে। আর আউটপুট যদি < **০.৫** হয়, তবে সেটিকে ক্লাস 0 হিসেবে গণ্য করে।

---

* বিশেষ দ্রষ্টব্য: লেকচারের শুরুতে বোর্ডে ব্র্যাকেটে "*cost function*" লেখা থাকলেও, এখানে মূলত মডেলের ***Hypothesis / Prediction Function*** (অনুমিত সিদ্ধান্ত সূত্র) আলোচনা করা হয়েছে। এর পরবর্তী ধাপে মডেলের ত্রুটি বা নির্ভুলতা পরিমাপের জন্য লজিস্টিক রিগ্রেশনের আসল কস্ট ফাংশন (যা *Binary Cross-Entropy* বা *Log Loss* নামে পরিচিত) শেখানো হয়ে থাকে।