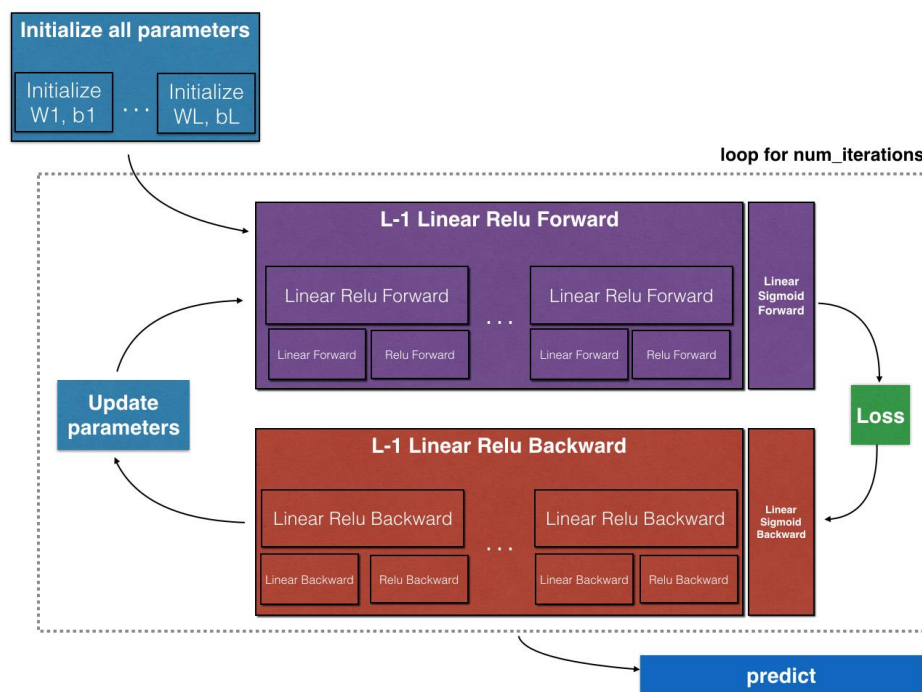


## پروژه سوم – ایجاد شبکه عصبی قدم به قدم

در این پروژه با هدف ایجاد شبکه عصبی قدم به قدم مراحل زیر را پیش رو داریم، تمامی مراحل ابتدا برای دو لایه و در مرحله بعد برای یک شبکه عمیق با  $L$  لایه در کد تکمیل کنید.



### مرحله صفر: اضافه کردن کتابخانه های مورد نیاز

فایل های موجود در پوشه Requirements را در مسیر پروژه قرار دهید.

### مرحله اول: مقدار دهی اولیه

مقدار دهی اولیه وزن ها و بایاس ها طبق آنچه در پروژه های پیشین تمرین کردید انجام میشود.  
(با کمک گرفتن از تابع `np.random.randn()*0.01`)

ساختار کلی شبکه به صورت زیر است:

LINEAR -> RELU -> LINEAR -> SIGMOID

### مرحله دوم: انتشار رو به جلو

بخش خطی ( $Z = Wx + b$ ) مرحله انتشار رو به جلو (با خروجی  $Z[1]$ ) را تکمیل کنید. توابع فعال ساز ReLU و سیگموید به شما داده شده است.

دو مرحله خطی و توابع فعال ساز را به صورت پیش سر هم در انتشار رو به جلو پیاده سازی کنید.

برای لایه اول تا L-1 از ترکیب خطی و فعالساز ReLU استفاده کنید.

برای لایه آخر تابع فعال ساز سیگموئید را کنار ترکیب خطی جایگزین کنید.

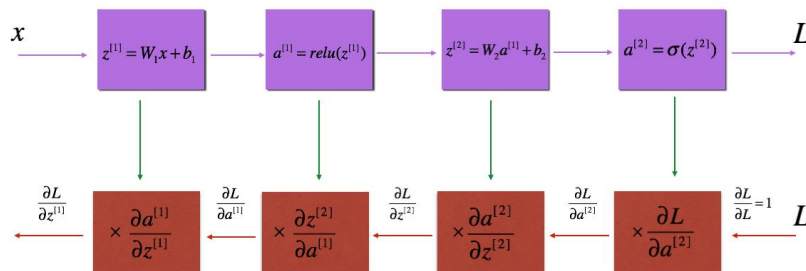
در نهایت تابع L\_model\_forward تکمیل میشود.

### مرحله سوم: محاسبه خطا

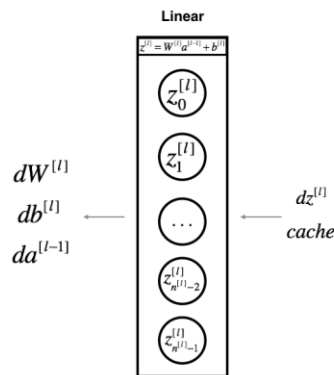
خطای شبکه از رابطه زیر محاسبه میشود:

$$-\frac{1}{m} \sum_{i=1}^m (y^{(i)} \log(a^{[L](i)}) + (1 - y^{(i)}) \log(1 - a^{[L](i)}))$$

### مرحله چهارم: پس انتشار



برای مرحله پس انتشار بخش خطی مناسب را پیاده سازی و در کد تکمیل کنید.



$$dW^{[l]} = \frac{\partial \mathcal{J}}{\partial W^{[l]}} = \frac{1}{m} dZ^{[l]} A^{[l-1]T}$$

$$db^{[l]} = \frac{\partial \mathcal{J}}{\partial b^{[l]}} = \frac{1}{m} \sum_{i=1}^m dZ^{[l](i)}$$

$$dA^{[l-1]} = \frac{\partial \mathcal{L}}{\partial A^{[l-1]}} = W^{[l]T} dZ^{[l]}$$

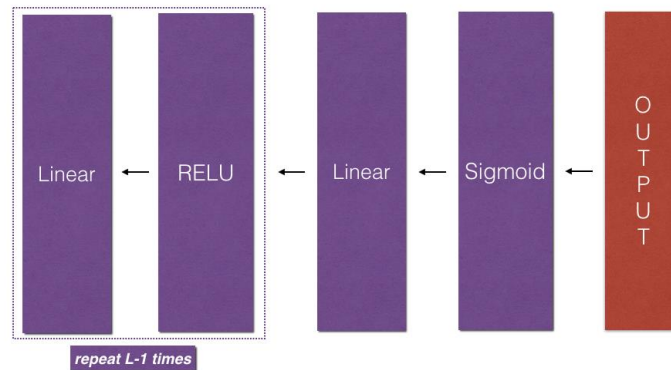
مشتق توابع فعال ساز به شما داده شده است. (در بخشهای Relu\_backward و Sigmoid\_backward)

مرحله خطی و مشتق تابع فعال ساز را ترکیب کنید.

$$dZ^{[l]} = dA^{[l]} * g'(Z^{[l]})$$

برای لایه L-1 تا لایه اول از Relu\_backward و برای لایه آخر از Sigmoid\_backward در کنار بخش خطی استفاده کنید.

در نهایت تابع L\_model\_backward تکمیل خواهد شد.



تابعی که در این بخش به آن نیاز دارید به صورت زیر است:

$$grads[" dW " + str(l)] = dW^{[l]}$$

مرحله پنجم: به روز رسانی وزن ها

در نهایت به روز رسانی وزن ها با توجه به نرخ یادگیری شبکه به صورت زیر است:

$$W^{[l]} = W^{[l]} - \alpha dW^{[l]}$$
$$b^{[l]} = b^{[l]} - \alpha db^{[l]}$$

روابط را در کد تکمیل کنید.