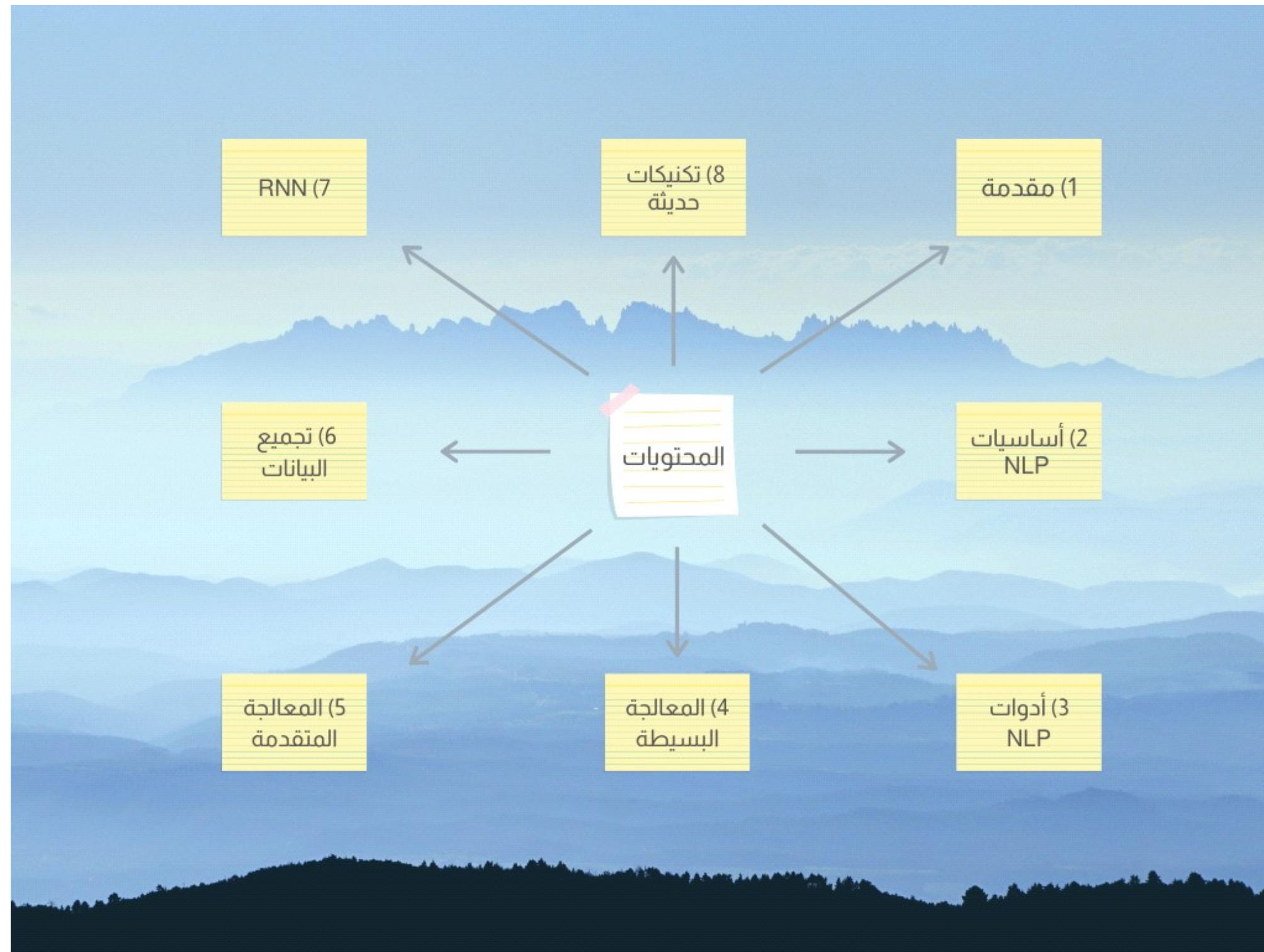


NATURAL LANGUAGE PROCESSING

المعالجة اللغوية الطبيعية



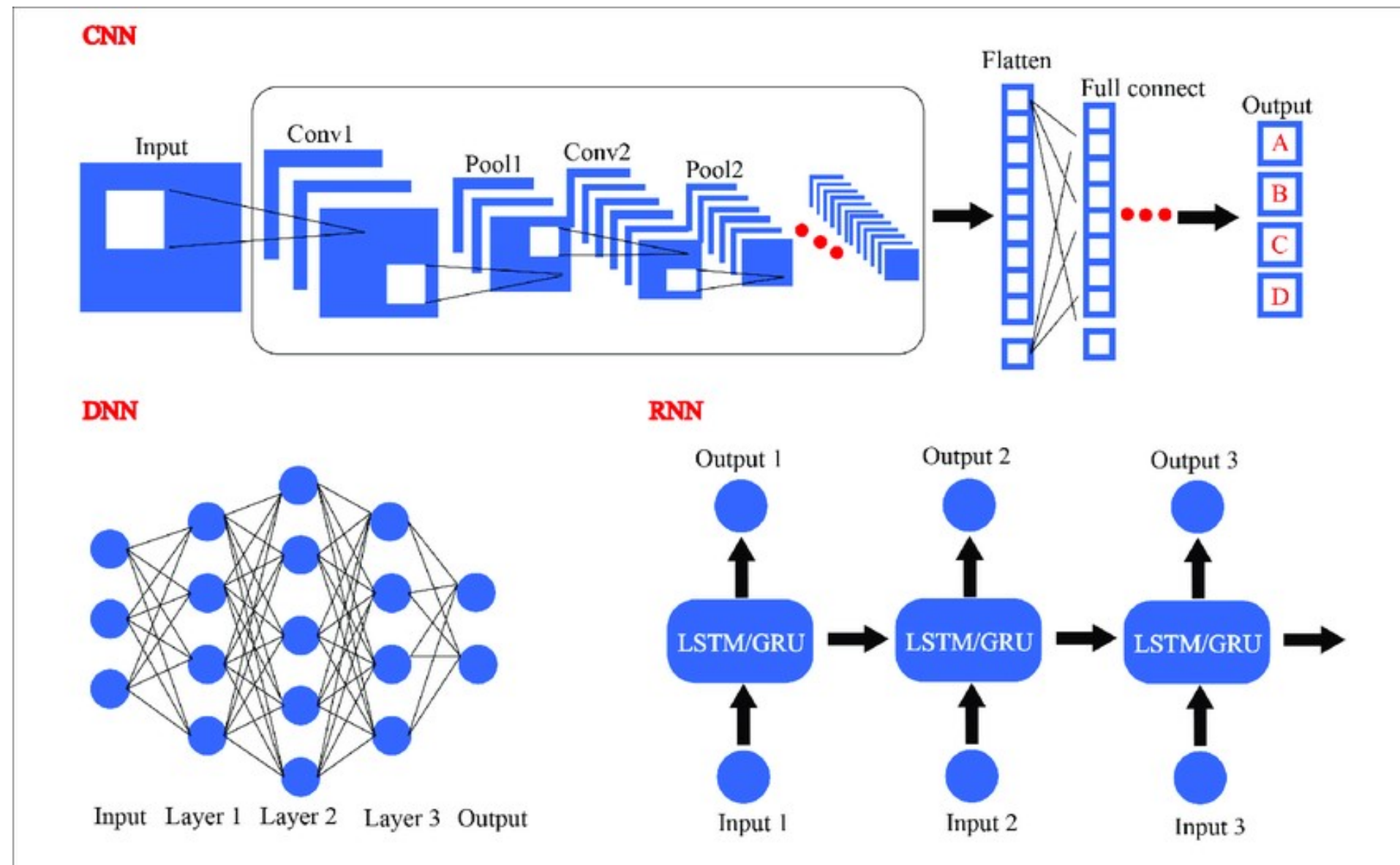
المحتويات

				التطبيقات	العقبات و التحديات	تاريخ NLP	ما هو NLP	المحتويات	(1) مقدمة
					البحث في النصوص	ملفات pdf	الملفات النصية	المكتبات	(2) أساسيات NLP
T.Visualization	Syntactic Struc.	Matchers	Stopwords	NER	Stem & Lemm	POS	Sent. Segm.	Tokenization	(3) أدوات NLP
	Dist. Similarity	Text Similarity	TF-IDF	BOW	Word2Vec	T. Vectors	Word embed	Word Meaning	(4) المعالجة البسيطة
T. Generation	L. Modeling	NGrams	Lexicons	GloVe	NMF	LDA	T. Clustering	T. Classification	(5) المعالجة المتقدمة
	Summarization & Snippets		Ans. Questions		Auto Correct	Vader	Naïve Bayes	Sent. Analysis	
Search Engine	Relative Extraction		Information Retrieval		Information Extraction		Data Scraping	Tweet Collecting	(6) تجميع البيانات
					Rec NN\TNN	GRU	LSTM	RNN	(7) RNN
Chat Bot	Gensim	FastText	Bert	Transformer	Attention Model	T. Forcing	CNN	Word Cloud	(8) تكتيكات حديثة

القسم السابع : الشبكات العصبية المتكررة

الجزء الأول : RNN

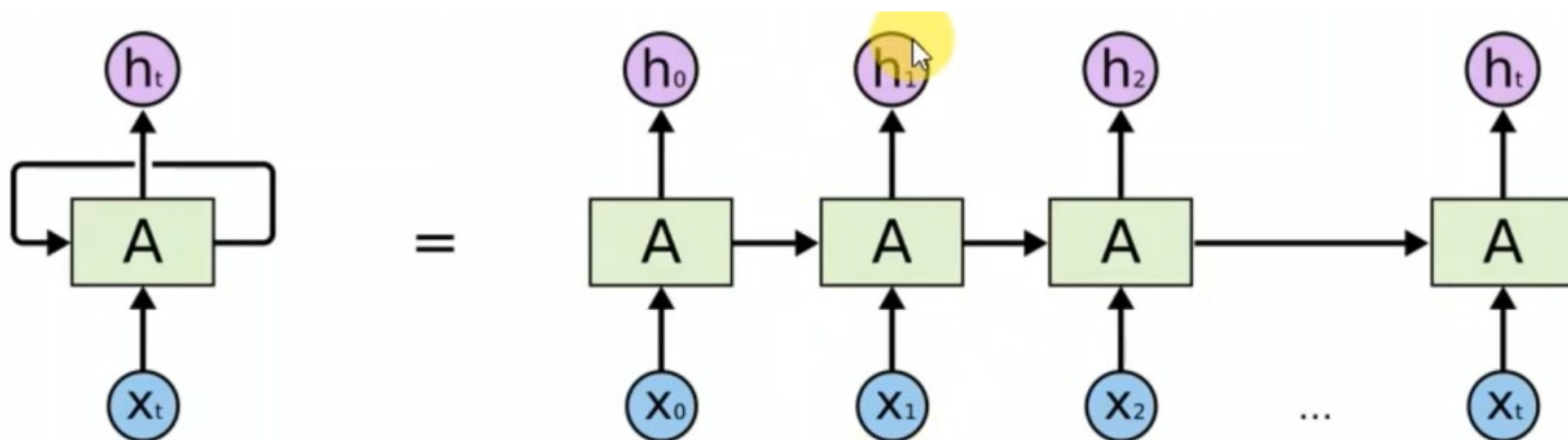
أنواع الشبكات العميقة : DNN , CNN , RNN



مفهوم ال-RNN:

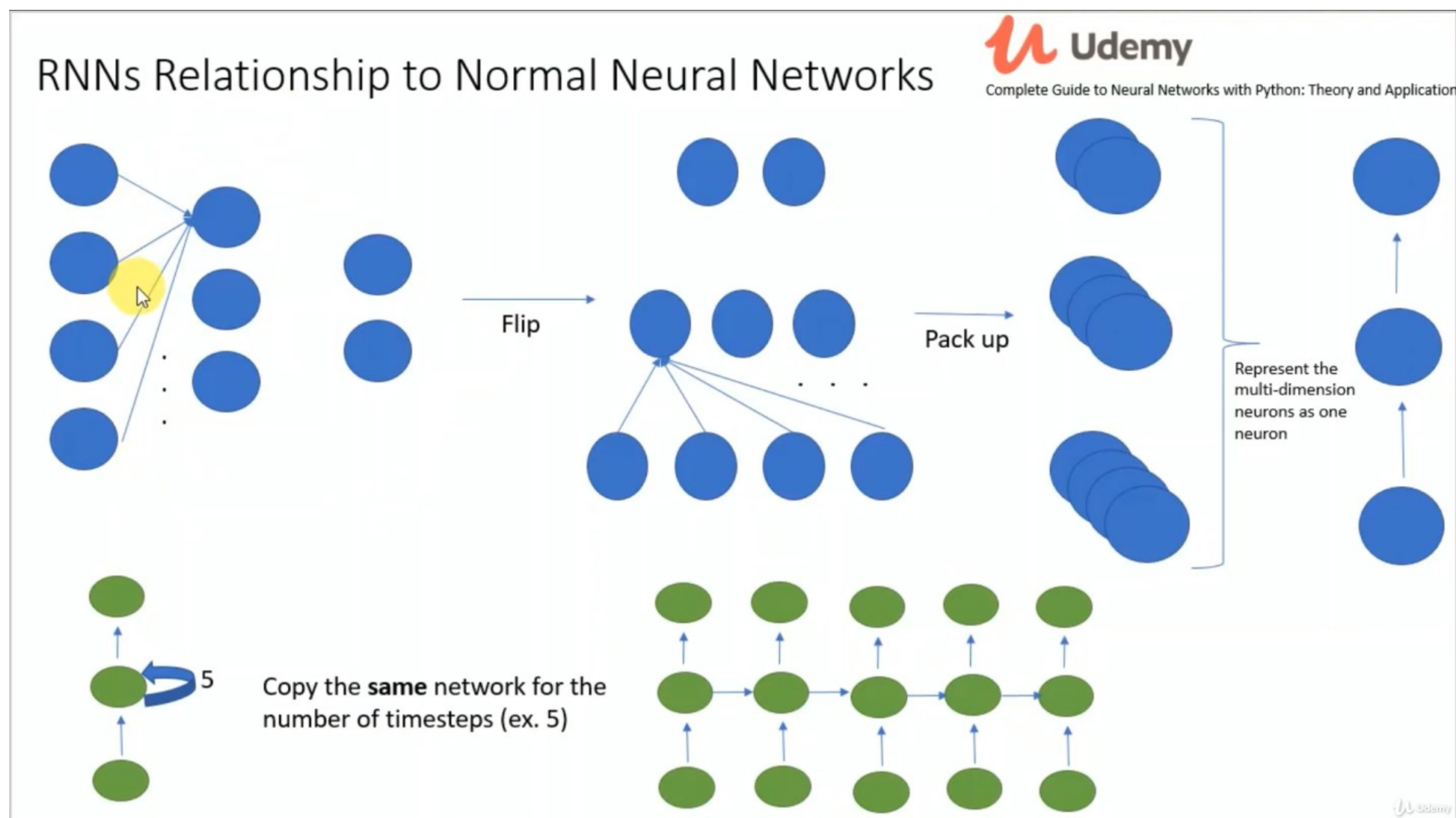
هي تستخدم بالأساس مع البيانات المتتابة sequence data , وهي المعتمدة علي عنصر الزمن , مثل اسعار الاسهم او بيانات الطقس و كذلك النصوص , لان جملة she spoke to her husband فكلمة her معتمدة علي كلمة she و كأن ما يميزها انها تتعامل كأن لديها ذاكرة تعتمد عليها

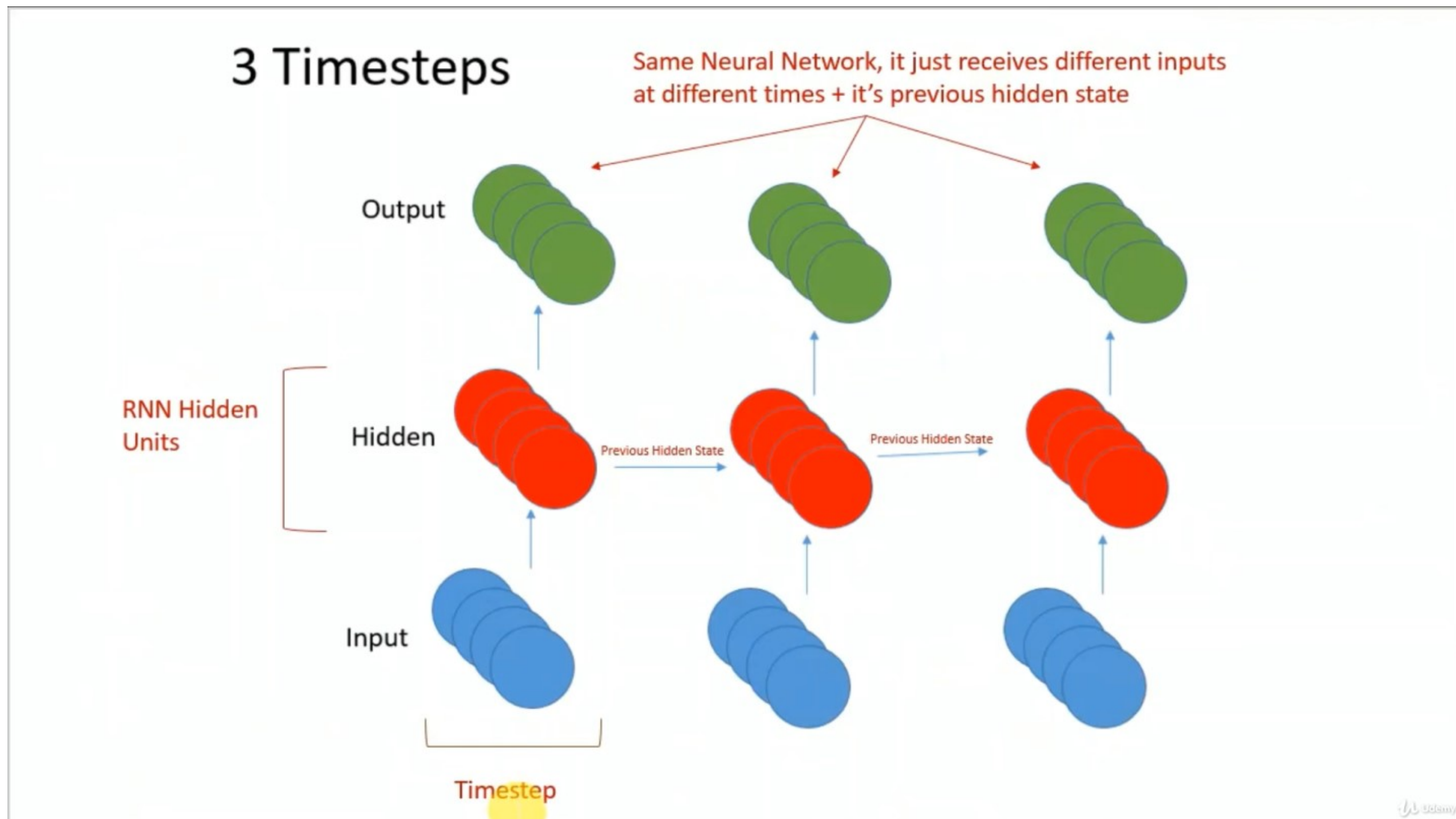
و تكون هيكلها كالتالي :



حيث تقوم الشبكة باستخدام المخرج علي انه مدخل مرة اخري

و تصميم RNN له علاقة بتصميم DNN , لان DNN , لو قمنا بتدويرها , ثم ضغطنا الخلايا , ثم كررناها , ستتحول الي RNN





مفهوم ال- Sequence to Sequence :

يطلق عليه أيضا (Seq2Seq) , ويقصد به الخوارزميات التي يكون لها المدخل عبارة عن بيانات متتابة , والمخرج بيانات متتابة كذلك

. والبيانات المتتابة هي البيانات المرتبطة بالعامل الزمني , مثل النصوص والصوت و الفيديو , لذا فهناك العديد من تطبيقات Seq2Seq في NLP مثل :

Translation , Abstracting, Text Generation , Chatbot , Question Answering

