

## **АНАЛИЗ СУЩЕСТВУЮЩИХ АРХИТЕКТУР НЕЙРОННЫХ СЕТЕЙ ДЛЯ ГЕНЕРАЦИИ ТЕКСТОВ ЕСТЕСТВЕННОГО ЯЗЫКА С ЦЕЛЮ ИССЛЕДОВАНИЯ АКТУАЛЬНЫХ ТЕХНИК ПРИ СОЗДАНИИ МОДЕЛЕЙ НЕЙРОННЫХ СЕТЕЙ**

**Полторак А.В., Набатов С.И.**

*МИРЭА - Российский технологический университет, 119454, Россия, г. Москва, проспект Вернадского, 78, e-mail: poltorak@mirea.ru, nabatov.s.i@mirea.ru*

---

**В данной статье рассмотрены существующие модели нейронных сетей, предназначенные для генерации естественного языка. Представлены их архитектуры, начиная от классической Seq2Seq модели, заканчивая генерацией естественного языка с помощью генеративно-сопоставительной сети, приведены различия между ними, проанализированы их преимущества и недостатки с точки зрения их точности при генерации естественного языка.**

---

Ключевые слова: нейронная сеть, обучение, модель нейронной сети, машинный перевод, Seq2Seq, GAN, генерация естественного языка, VAE, эмбединги.

## **ANALYSIS OF EXSITING NEURAL NETWORK ARCHITECTURES FOR THE GENERATION OF NATURAL LANGUAGE TEXTS FOR THE PURPOSE OF RESEARCHING CURRENT TECHNIQUES IN CREATING NEURAL NETWORK MODELS**

**Poltorak A.V., Nabatov S.I.**

*MIREA - Russian Technological University, 119454, Moscow, 78 Vernadskogo Avenue, Russia, e-mail: poltorak@mirea.ru, nabatov.s.i@mirea.ru*

---

**This article discusses the existing models of neural networks designed to natural language generation. Their architectures are presented, ranging from the classical Seq2Seq model to the generation of a natural language using a generative adversarial network, the differences between them are given, their advantages and disadvantages are analyzed in terms of their accuracy of generating a natural language.**

---

Keywords: neural network, learning, neural network model, machine translation, Seq2Seq, GAN, natural language generation, VAE, embeddings.

### **Введение**

Точкой отправления нейронных сетей для генерации текстов является рекуррентная нейронная сеть (англ. RNN). Рекуррентная нейронная сеть - это тип нейронных сетей, который использует направленную последовательность для связей между элементами [18]. Такая архитектура нейронной сети позволяет запоминать определенные части предложения, однако без учета смысла самих слов. Чтобы учесть смысл слов, были придуманы “вложения слов” (англ. Word embeddings), которые могут представлять слово в виде тензора (в частном случае вектора) [19]. Одни из архитектур RNN - это LSTM (Long-Short Term Memory) и GRU (Gated Recurrent Unit).

В этой статье рассмотрены существующие модели нейронных сетей, предназначенные для генерации естественного языка. Представлены их архитектуры, начиная от классической Seq2Seq модели, заканчивая генерацией естественного языка с помощью генеративно-сопоставительной сети, приведены различия между ними, проанализированы их преимущества и недостатки. В статьях [15 - 17] приведены основные архитектуры и техники для генерации текстов, однако в них не рассмотрены новые техники, что описаны в данной статье.

### “Sequence-to-sequence” - первая модель

“Sequence to sequence” - это архитектура, созданная компанией “Google” для решения задачи “машинного перевода”. В оригинальной статье слоями сети являлись RNN и LSTM [1], которые позволяют решить проблему градиентного взрыва, затем также было предложено использовать GRU слой, который является менее сложной версией LSTM [13]. Контекстом каждого входа, является выход с предыдущего момента времени.

Перед тем, как рассмотреть способ генерации естественного языка, необходимо рассмотреть, что такое энкодеры и декодеры в контексте нейронных сетей, поскольку “Seq2Seq” модель состоит из трех частей: энкодер (enc), промежуточный вектор (vec) и декодер (dec) [1].

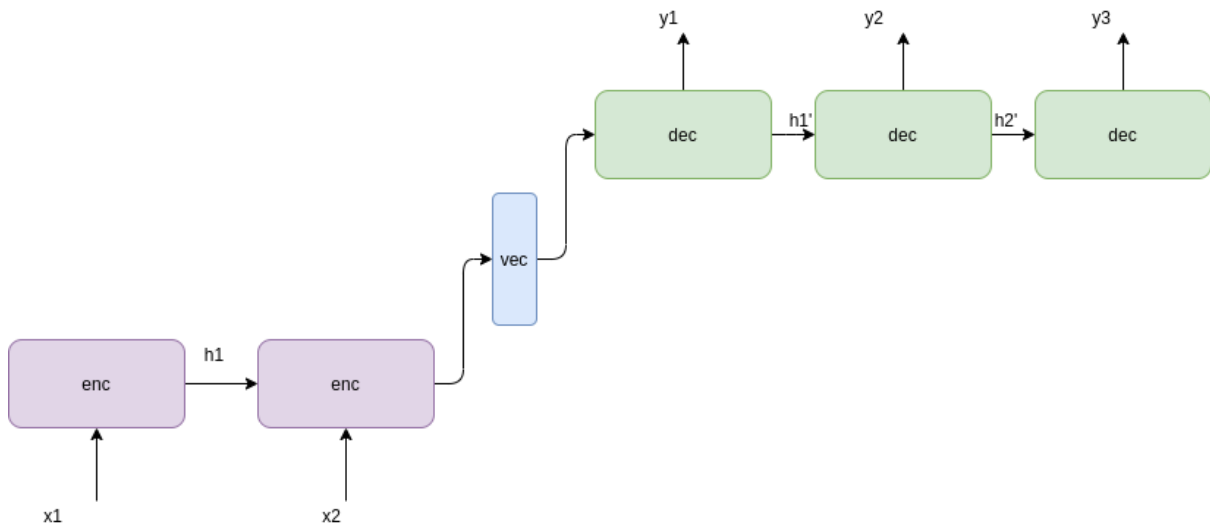


Рис. 1. Seq2Seq модель – Первая модель

Энкодер служит для того, чтобы получить на вход предложение переменной длины на исходном языке и перевести его в промежуточный вектор. Промежуточный вектор, в свою очередь, дает статическое представление закодированного предложения. После этого вектора закодированное предложение передается в декодер, который ставит ему в соответствие предложение другой длины. Таким образом, каждая ячейка - это слой нейронной сети, у которого есть свои веса.

Обычно “Seq2seq” модель используется для машинного перевода одного языка в другой. Но тут, интересно, где находится значащее пространство (*англ. meaning space*), а оно находится только внутри промежуточного вектора, который должен собрать представление предложения на одном языке и передать его декодеру.

Несмотря на то, что “Seq2seq”, благодаря своей гибкой архитектуре “энкодер-декодер”, нашел применение во множествах прикладных задач, помимо машинного перевода, например, перенос стилей языка [2], генерация описаний для картинок [3], задача ответа на визуальный вопрос [4] и многое другое, **в этой архитектуре есть “узкое место” - это промежуточный вектор**, поскольку его размер он статический, таким образом, модель сможет запоминать лишь предложения фиксированной длины. Рис. 2 отражает постепенную потерю навыков обобщения модели при возрастании длины предложения.

#### Значащее пространство

Значащее пространство - это представление и сопоставление **концептов и идей**. Если не нужно представление, а достаточно прикладного применения, то можно полагать, что это пространство является последним состоянием энкодера. Однако для генерации языка, или для применения дополнительных стилей к языку, необходимо расширить представления слов в этом пространстве. Также мы можем добавить определенное ограничение в форме вероятностного распределения, что позволит нам представить это пространство более компактно.

Если наша система будет способна сохранять этот баланс, заданный распределением, то мы можем утверждать, что значения, которые будут иметь похожее значение будут распределены близко друг к другу. На практике, это означает следующее: на выходе энкодера рекуррентной нейронной сети выход последнего состояния передается двум линейным слоям, чтобы те конвертировали это состояние в

среднее значение и в ковариационный вектор [5], такой подход назван вариационным методом. На рис. 3 представлен вариационный автоэнкодер.

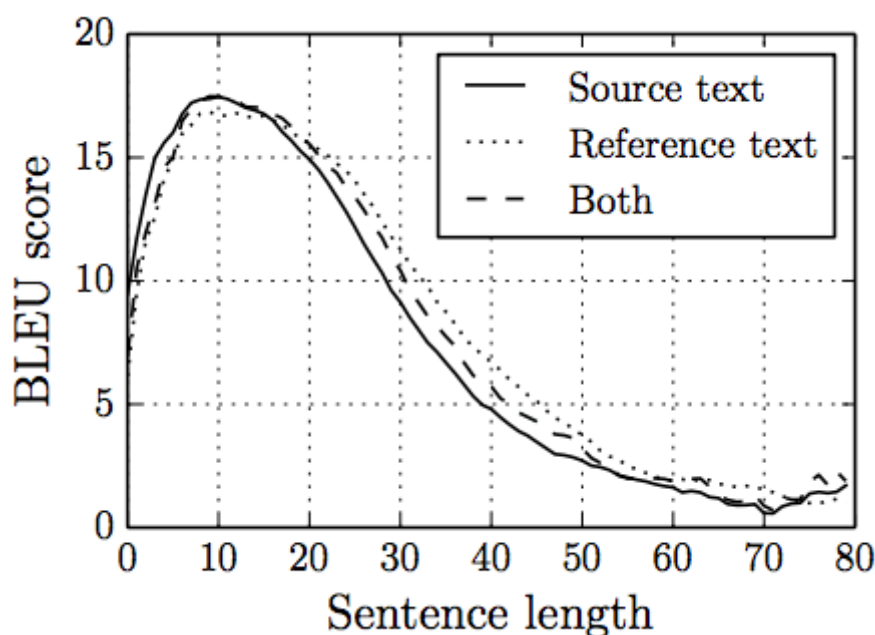


Рис. 2. Зависимость длины предложения от BLEU [14]

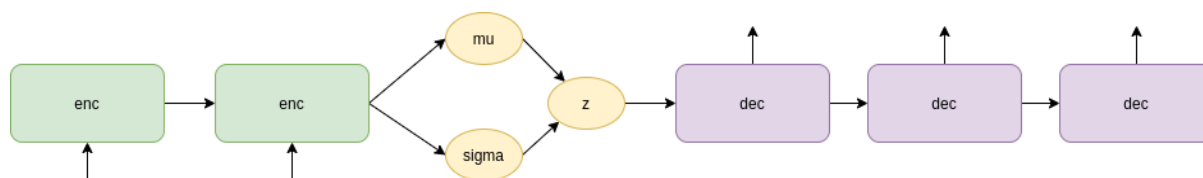


Рис. 3. VAE – вторая модель

На вход подается текст  $x$ , который мы кодируем в некое значимое пространство, которое представляет собой распределение вероятностей  $q(z|x)$ . Вариационный автоэнкодер заменяет стандартную детерминированную функцию, которая была на конце автоэнкодера, на обученную модель  $q(z|x)$ , которая обычно реализует приближенное Гауссово распределение по  $z$ . Введем  $p(z)$  - это будет идеальное Гауссово распределение, которое обозначается как  $N(0, 1)$ , следовательно, наша модель  $q(z|x)$  должна стремиться быть похожей на  $p(z)$ .

После того, как  $x$  был закодирован в распределение  $q(z|x)$ , передается  $z$  для того, чтобы восстановить  $x$ . Проблема VAE заключается в том, что его тяжело тренировать, поддерживая корректное распределение.

Поскольку внутри значащего пространства теперь распределение, можно поставить внутри распределения функцию по определенным параметрам, преобразовывая по определенным параметрам, т.е. в частном случае искать синонимы, например, преобразование современного текста в шекспировскую речь [6].

#### Избавление от парных данных

Несмотря на плюсы и минусы Seq2Seq моделей, рассмотренных ранее, в них есть одна общая черта - это необходимость использования парных данных при обучении, например, английского языка с русским.

Одно из решений этой проблемы предложил Google, например, имеется две пары языков - (Английский, Корейский) и (Английский, Японский), необходимо транслировать из корейского в японский, и наоборот. В их систему перевода был добавлен токен, показывающий в какой язык надо переводить, например <2 EN>, что означает “перевод в английский”. Также, они заметили, что, если сопоставить в значащем пространстве несколько языков, то сеть “делится” параметрами между представленными языками, следовательно, система позволяет сопоставлять слова между парами языков.

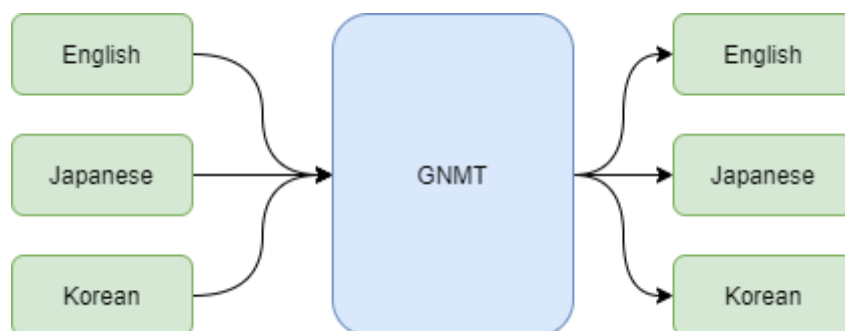


Рис. 4. GNMT – третья модель

В случае если упорядоченных пар нет, можно использовать общий энкодер у этой пары языков (A, B), но значащее пространство тренируется как автоэнкодер для воспроизведения различных языков (рис. 5) [8]. Например, чтобы превратить язык A в язык B, можно энкодировать предложения в язык B и затем декодировать их в язык A через общее пространство, таким образом, мы получим эту пару (A, B). Это же правило действует и в обратную сторону.

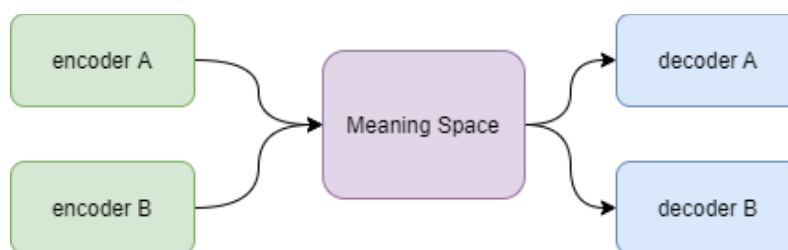


Рис. 5. Использование общего значащего пространства для создания упорядоченных пар языков

### Состязательная генерация текста

Генеративно-состязательные сети [9] также используются для генерации текста без существующих пар предложений посредством классификации с учителем, заключающейся в отличии сгенерированного текста от требуемого текста.

Следующий метод [10] фактически описывает вариацию классической статьи про генеративно-состязательные сети, только заменяя генератор на Seq2Seq модель со значащим пространством, которое распределено близко к нормальному распределению  $N(0, 1)$ . Выход этого генератора проходит через классификатор, также называемый дискриминатором. Этот классификатор уже обучен необходимой генерации, т.е. он знает, что должно быть на выходе (рис. 6).

Но в таком подходе возникает следующая проблема: текст - это дискретная величина, а не продолжительная, как в случае с изображениями, вследствие чего мы не можем дифференцировать по этому пространству и не можем сделать метод обратного распространения ошибки. В модели на рис. 5 рассмотрен один из вариантов решения этой проблемы - это использование среднего взвешенного всех слов векторов.

Еще один подход называется “professor forcing” [11 - 12]. Этот подход отличается тем, что дискриминатор смотрит не на выход декодера, а на разницу между скрытыми состояниями, в то время как декодер обучается в режиме принуждения учителя (*англ.* - “teacher forcing”). То есть мы хотим, чтобы поведение генеративной рекуррентной нейронной сети было неотличимым вне зависимости от режима тренировки, будь то определенная последовательность, или это самогенерированные входы.

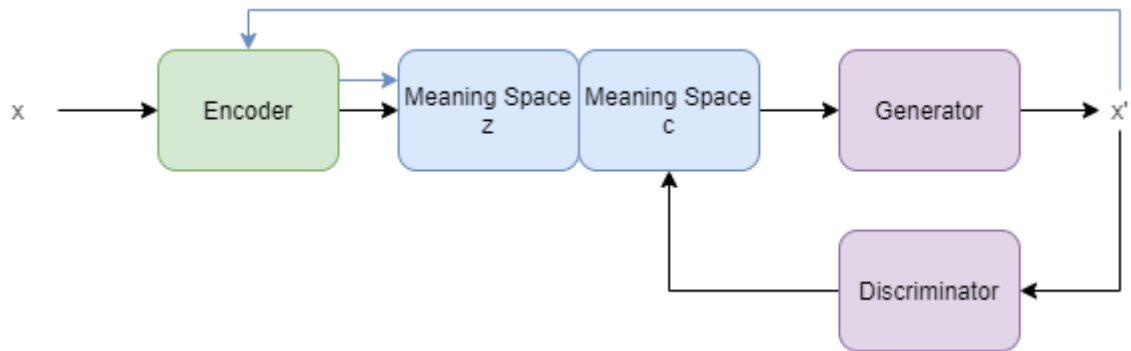


Рис. 6. Генеративная модель – четвертая модель

Таким образом, параметры дискриминатора тренируются, чтобы максимизировать правдоподобие, что достигается с помощью стохастического градиентного спуска с мини батчами, которые представляют собой комбинацию  $N$  последовательностей, которые были получены при “teacher forcing”, и  $N$  последовательностей, полученных в режиме генерации. Параметры же рекуррентной нейронной сети тренируются, чтобы максимизировать правдоподобие и обмануть дискриминатор.

На рис. 7 представлена архитектура “Professor Learning”. Распределения скрытых состояний принудительно заставляют быть близкими друг к другу благодаря дискриминатору.

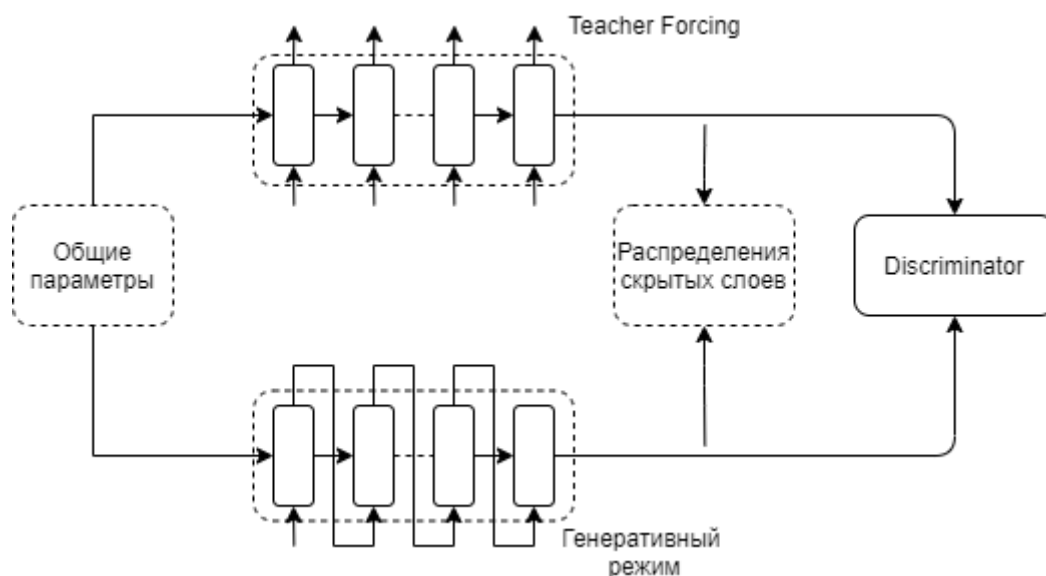


Рис. 7. Архитектура “Professor Learning”

### Заключение

В рассмотренных моделях нейронных сетей есть несколько общих проблем, например, сложность обучения у всех генеративно-состязательных моделей, поскольку помимо стандартной для таких типов сетей тренировки двух моделей одновременно, добавляется еще необходимость следить за распределениями скрытых слоев у двух частей генератора при архитектуре “Professor Learning”, также текст - это дискретная величина и потому, необходимо каким-то образом обновлять веса методом обратного распространения ошибки, что также осложняет обучение.

С другой стороны, когда используется уже классическое обучение рекуррентных нейронных сетей, необходимо собирать парный датасет, что во многих случаях сделать тяжело, в связи с нехваткой данных.

Таким образом, главной текущей задачей для моделей нейронных сетей для генерации естественного языка является создание таких моделей, которые позволяют работать с непарными датасетами, например, GNMT, GAN с “Professor Learning” и, в целом, различные вариации использования значащего пространства.

1. Ilya Sutskever, Oriol Vinyals, Quoc V. Le. Sequence to Sequence Learning with Neural Networks // NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2 December 2014 - P. 3104–3112
2. Harsh Jhamtani, Varun Gangal, Eduard Hovy, Eric Nyberg. Shakespearizing modern language using copy-enriched sequence-to-sequence models. Proceedings of the Workshop on Stylistic Variation // EMNLP - 2017 – P. 10–19.
3. Andrej Karpathy, Li Fei-Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions // CVPR - 2015 – Paper Conference
4. Bashar Talafha Mawdoo3 AI Mahmoud Al-Ayyoub. JUST at VQA-Med: A VGG-Seq2Seq Model // Conference: CLEF Avignon - July 2018 – Paper Conference.
5. Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew M. Dai, Rafal Jozefowicz & Samy Bengio. Generating Sentences from a Continuous Space // SIGNLL Conference on Computational Natural Language Learning (CONLL) – 2016 – Paper Conference.
6. Jonas Mueller, David Gifford, Tommi Jaakkola. Sequence to Better Sequence: Continuous Revision of Combinatorial Structures: Proceedings of the 34th International Conference on Machine Learning // PMLR 70 - 2017 – P. 2536-2544
7. Mike Schuster, Melvin Johnson and Nikhil Thorat. Zero-Shot Translation with Google's Multilingual Neural Machine Translation System // Google: - November 22, 2016 – Paper Conference:
8. Mikel Artetxe, Gorka Labaka & Eneko Agirre. Unsupervised Neural Machine Translation // ICLR – 2018 – Paper Conference.
9. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative adversarial nets: Advances in Neural Information Processing Systems // Curran Associates, Inc., 27 – 2014 – P. 2672 - 2680.
10. Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, Eric P. Xing. Toward Controlled Generation of Text // ICML - 2017 – Paper Conference
11. Alex Lamb, Anirudh Goyal, Ying Zhang, Saizheng Zhang, Aaron Courville, Yoshua Bengio. Professor Forcing: A New Algorithm for Training Recurrent Networks // NIPS - 2016 – Paper Conference.
12. Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu , Tie-Yan Liu, Wei-Ying Ma. Dual Learning for Machine Translation // NIPS – 2016 – Paper Conference.
13. Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling // NIPS – 2014 – Paper Conference.
14. Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, Yoshua Bengio. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches // Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, October – 2014 – P. 103 - 111.
15. Cristina Garbacea, Qiaozhu Mei. Neural Language Generation: Formulation, Methods, and Evaluation // CoRR – 2020 – Paper Conference.
16. Sashank Santhanam, Samira Shaikh. A Survey of Natural Language Generation Techniques with a Focus on Dialogue Systems - Past, Present and Future Directions // CoRR – 2019 – Paper Conference.
17. Dupond, Samuel. A thorough review on the current advance of neural network structures // Annual Reviews in Control, 14 – 2019 – P. 200–230.
18. Tealab, Ahmed. Time series forecasting using artificial neural networks methodologies // A systematic review. Future Computing and Informatics Journal. 3 (2) – 2018 – P. 334–340.
19. Jurafsky, Daniel; H. James, Martin. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, N.J.: Prentice Hall. ISBN 978-0-13-095069-7.

1. Ilya Sutskever, Oriol Vinyals, Quoc V. Le. Sequence to Sequence Learning with Neural Networks // NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2 December 2014 - P. 3104–3112
2. Harsh Jhamtani, Varun Gangal, Eduard Hovy, Eric Nyberg. Shakespearizing modern language using copy-enriched sequence-to-sequence models. Proceedings of the Workshop on Stylistic Variation // EMNLP - 2017 – P. 10–19.
3. Andrej Karpathy, Li Fei-Fei. Deep Visual-Semantic Alignments for Generating Image Descriptions // CVPR - 2015 – Paper Conference
4. Bashar Talafha Mawdoo3 AI Mahmoud Al-Ayyoub. JUST at VQA-Med: A VGG-Seq2Seq Model // Conference: CLEF Avignon - July 2018 – Paper Conference.
5. Samuel R. Bowman, Luke Vilnis, Oriol Vinyals, Andrew M. Dai, Rafal Jozefowicz & Samy Bengio. Generating Sentences from a Continuous Space // SIGNLL Conference on Computational Natural Language Learning (CONLL) – 2016 – Paper Conference.
6. Jonas Mueller, David Gifford, Tommi Jaakkola. Sequence to Better Sequence: Continuous Revision of Combinatorial Structures: Proceedings of the 34th International Conference on Machine Learning // PMLR 70 - 2017 – P. 2536-2544
7. Mike Schuster, Melvin Johnson and Nikhil Thorat. Zero-Shot Translation with Google's Multilingual Neural Machine Translation System // Google: - November 22, 2016 – Paper Conference:
8. Mikel Artetxe, Gorka Labaka & Eneko Agirre. Unsupervised Neural Machine Translation // ICLR – 2018 – Paper Conference.
9. Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. Generative adversarial nets: Advances in Neural Information Processing Systems // Curran Associates, Inc., 27 – 2014 – P. 2672 - 2680.
10. Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, Eric P. Xing. Toward Controlled Generation of Text // ICML - 2017 – Paper Conference
11. Alex Lamb, Anirudh Goyal, Ying Zhang, Saizheng Zhang, Aaron Courville, Yoshua Bengio. Professor Forcing: A New Algorithm for Training Recurrent Networks // NIPS - 2016 – Paper Conference.
12. Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu , Tie-Yan Liu, Wei-Ying Ma. Dual Learning for Machine Translation // NIPS – 2016 – Paper Conference.
13. Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling // NIPS – 2014 – Paper Conference.
14. Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, Yoshua Bengio. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches // Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation, October – 2014 – P. 103 - 111.
15. Cristina Garbacea, Qiaozhu Mei. Neural Language Generation: Formulation, Methods, and Evaluation // CoRR – 2020 – Paper Conference.
16. Sashank Santhanam, Samira Shaikh. A Survey of Natural Language Generation Techniques with a Focus on Dialogue Systems - Past, Present and Future Directions // CoRR – 2019 – Paper Conference.
17. Dupond, Samuel. A thorough review on the current advance of neural network structures // Annual Reviews in Control, 14 – 2019 – P. 200–230.
18. Tealab, Ahmed. Time series forecasting using artificial neural networks methodologies // A systematic review. Future Computing and Informatics Journal. 3 (2) – 2018 – P. 334–340.
19. Jurafsky, Daniel; H. James, Martin. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, N.J.: Prentice Hall. ISBN 978-0-13-095069-7.