

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
1	Text Generation	An Integrated Deep Generative Model for Text Classification and Generation	ZhengWang y QingbiaoWu	2018	Auto encoders, Generative model, Text generations, Text classification, Natural Language Processing Systems	Bag of words, Word embedding, Text label	Variational Autoencoder (VAE), LSTM	En este artículo, se propone un modelo integrado basado en VAE para manejar tareas de generación y clasificación de texto.	3	Sirve también para la generación de texto, utiliza explícitamente la información de la etiqueta durante la decodificación, resultados competitivos en ambas tareas	El modelo fue entrenado para la clasificación de texto	Resultados de clasificación competitivos y el texto generado es realista	Introducir algunos mecanismos nuevos para mejorar el decodificador, como la probabilidad de cobertura. Introducir algunas funciones de restricción para producir un texto más controlable después de modificar la función objetivo.
2	Text Generation	Recent advances of neural text generation- Core tasks, datasets, models and challenges	JIN HanQi, CAO Yue, WANG TianMing, XING XinYu & WAN XiaoJun	2020	Natural language generation (NLG), neural text generation, AMR-to-text, data-to text, text summarization, paraphrase generation	No se especifican técnicas de PLN	Métricas de evaluación: Perplexity, BLEU (based n-gram), meteor, Self-BLEU Modelos: Seq2Seq, Transformers, GPT-2, BERT, VAE, GAN	Esta review tiene como objetivo proporcionar una síntesis actualizada de las tareas centrales en la generación de texto neuronal y las arquitecturas adoptadas para manejar estas tareas, y sobre los desafíos en la generación de texto neuronal.	2	Presenta los avances recientes en la generación de texto neuronal. - Guía y referencia para investigadores y profesionales en esta área.	No entran en detalle, ni mencionan técnicas específicas de PLN para la generación de texto	-	No se especifica el trabajo futuro
3	Text Generation	LSTM encoder-decoder with adversarial network for text generation from keyword	Dongju Park, Chang Wook Ahn	2018	Text generation, Generative Adversarial Network, Natural Language Processing	Word Embedding, Skip gram, Word2Vec	NLG, GAN, RNN, LSTM, selfAttention, MLE, BLEU Score, COCO	En este artículo, se propone un modelo para generar texto a partir de una palabra determinada. El propósito es contribuir a ampliar la diversidad de textos generados a partir de una sola palabra.	4	Generación de texto realista basado en una palabra dada y un discriminador que puede distinguir entre datos generados y reales Uso de LSTM en lugar de RNN en el codificador y decodificador Mayor diversidad de oraciones	- Genera oraciones cortas	- El modelo obtiene buen BLEU score en comparación a otros modelos que también son buenos	Expandir el modelo para que la relación entre las palabras de entrada se pueda ingresar juntas, o el modelo se pueda inferir directamente para que funcione bien para generar varias oraciones.
4	Text Generation	A Text Generation and Prediction System: Pretraining on New Corpora Using BERT and GPT-2	Yuanbin Qu, Peihan Liu, Wei Song, Lizhen Liu, Miaomiao Cheng	2020	Language Model, Text Generation, OpenAI GPT-2, BERT	No se especifican técnicas de PLN	GPT-2, BERT, Transformer	En este documento, se utilizan modelos pre-entrenados para completar algunas tareas de generación de texto chino, incluida la generación de oraciones largas, utilizando Transformer a partir de modelos pre entrenados	4	Oraciones largas a partir de palabra clave Desempeño de GTP-2 de vanguardia Se describe el pseudocódigo del modelo de generación de texto a grandes rasgos	Para oraciones demasiadas largas comienza a volverse repetitivo (duplicaciones), se aclara que el problema se puede deber a que los corpus seleccionados no están estandarizados, por lo que el modelo no logra aprender todas las reglas del lenguaje. Entrenado para generar texto en chino	Oraciones largas que generalmente, el significado y la legibilidad de la mayoría de las oraciones generadas son apropiados, además, consistente con la palabra de entrada	Utilizar el modelo para realizar más tareas Continuar procesando los corpus de acuerdo con el modelo existente Intentar mejorar los contras de su modelo
5	Question Answering	Automated Query Analysis Techniques for Semantics based Question Answering System	Shrimai Prabhume, Piyush Rai, Loverose S. Sandhu, Priya L, Sowmya Kamath S	2014	Question answer systems, Natural language understanding (NLU), Query processing, Text analysis	Relaciones semánticas, Etiquetado de parte del discurso, Análisis sintáctico	WordNet	En este artículo se presenta un prototipo de motor de búsqueda de preguntas y respuestas que utiliza PLN, NLG, clasificación de preguntas y registros de consultas para encontrar una respuesta precisa a la consulta enviada	3	El sistema propuesto integra varias técnicas de optimización, como registros de consultas y clasificación de preguntas.	Trabajo enfocado a motores de búsqueda	-	Mejorar el análisis semántico para poder analizar más las consultas complejas Incluir más parámetros en el análisis semántico
6	Automatic Generation Questions	Automatic question generation	Mark Last, Guy Danon	2020	Text mining, Automatic Generation, Question answering systems, NLP, Search engines	Semantic Role Labeling, Named Entity Recognition	ARK, SQuAD, WordNet	Esta revisión trata sobre generación automática de preguntas, describen los pasos principales en la construcción, clasificación y evaluación de preguntas; además, discuten sobre los desafíos de AQG en el campo del NLP	3	Muchas referencias explicadas sobre trabajos realizados en el campo AQG - NLP de otros autores y sus logros	La aplicación de sistemas AQG enfrenta múltiples desafíos en la actualidad. AQG en dominios multilingües y multilingües necesitan una mayor exploración Se desmarca un poco del campo de NLP a pesar de pertenecer a este campo (Requiere conocimientos adicionales a NLP)	-	No se especifica el trabajo futuro

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
7	Text Generation	Text Generation Using Long Short-Term Memory Networks	Ishika Dhall, Shubham Vashisth and Shipra Saraswat	2020	LSTM networks, RNN, Natural Language Processing, Text Generation, Machine Learning	No se especifican técnicas de PLN	RNN, LSTM, Keras, Tensorflow, Adam	En este artículo se presenta una red de memoria a largo plazo, profunda y apilada, una forma avanzada de modelo de red neuronal recurrente que puede generar texto a partir de una semilla de entrada aleatoria; además, presenta las complicaciones de RNN	2	- Se especifica cuáles son los pasos a grandes rasgos que se deben seguir para construir un modelo con LSTM	No se propone ningún modelo, ni se muestra cómo crear o entrenar el modelo LSTM Puramente teórico Entrenar este tipo de modelos LSTM es computacionalmente costoso	Los datos generados por el modelo LSTM son de naturaleza realista y de naturaleza interpretable El modelo entrenado fue capaz de generar texto para una semilla generada aleatoriamente con una precisión de prueba del 71,22%	- Superar la precisión de los modelos LSTM actuales agregando más capas, nodos a la red y aplicando la noción de transferencia de aprendizaje en el mismo dominio de problemas
8	Question Answering	A System for Generating Cloze Test Items from RussianLanguage Text	Andrey Kurtasov	2013	NLP, Cloze question generation, Moodle LMS, Formal grammars	Sentence Splitting, Sentence Filtering, Summarization (extraction and abstraction), Question Generation	-	En este artículo se estudia el problema de generación de pruebas educativas automatizadas. Se Describe un procedimiento para generar elementos de prueba cloze a partir de texto en ruso.	4	Descripción detallada de serie de pasos para generar pruebas automatizadas	El artículo cuenta con varios años desde su publicación y las estrategias usadas parecen estar desfasadas, sin embargo, la metodología sigue vigente Los resultados que obtuvieron carecen de precisión a nivel de NLP Orientado al idioma ruso	Creación de un sistema que genera preguntas a partir de textos electrónicos con fin educativo y evaluativo en idioma ruso	Mejorar la calidad del sistema, la generación de preguntas y su coherencia
9	Question Answering	Automated Generation of Assessment Test Items from Text: Some Quality Aspects	Andrey Kurtasov	2013	NLP, Cloze question generation, Moodle LMS, e-learning	Bag of words, Sentence splitting	-	En este artículo se describe el problema de la generación automatizada de elementos de prueba de evaluación a partir de texto en lenguaje natural	3	Se mencionan las técnicas de NLP que podrían ser usadas para mejorar la calidad del modelo a desarrollar	-	-	Mejorar interfaz de usuario Usar herramientas de Summarization Resolución de anáforas, generación de oraciones interrogativas y generación de distractores para preguntas de opción múltiple
10	Text Generation	Recurrent Convolution Attention Model (RCAM) for Text Generation based on Title	Yuan Jianglin, Guo Zhigang and Chen Gang	2019	Convolution, Deep learning, Natural language processing systems, NLP, NLG, LSTM	Word2Vec	RCAM, LSTM, RNN, CNN, GAN, SeqGAN, TextRank, Tensorflow, BLEU	En este artículo se propone un modelo de atención de convolución recurrente con LSTM para generar texto a partir de un título, enfocado en la generación de texto a texto combinando ventajas de CNN y LSTM	3	Se comparan modelos entrenados con diferentes tipos de redes neuronales mostrando sus ventajas y desventajas para la tarea de generación de texto a partir de un título Se especifica el modelo mostrando gráficos, pseudocódigo y fórmulas	- Dependencias a largo plazo y errores de sintaxis en algunas partes del texto generado	Generación de oraciones de acuerdo con el título y hacer que el texto se exprese con mayor fluidez Menos tiempo para entrenar en contraste con otros tipos de redes neuronales	Corregir las dependencias a largo plazo y errores de sintaxis en el texto generado
11	Question Answering	Automatic Multiple Choice Question Generation From Text: A Survey	Dhawaleswar Rao and Sujana Kumar Saha	2020	Automatic question generation, multiple choice questions MCQ, natural language processing, text analysis	Text Normalization, Structural Analysis, Sentence Simplification, Lexical Analysis, Statistical Analysis, Syntactic Analysis, Coreference resolution, Word Sense Disambiguation (WSD)	WordNet, N-gram	En este documento se hace una revisión sistemática de sistemas de Multiple Choice Questions, presentando hallazgos sobre la revisión. Se describe el flujo de trabajo genérico para un sistema automático de generación de MCQ.	5	Se aborda cada paso llevado a cabo para la generación de preguntas de múltiple respuesta, además, se mencionan los trabajos que están enfocados en este ámbito, especialmente sistemas de MCQ	-	Establecimiento de flujo de trabajo genérico que consta de seis fases dependientes ampliamente clasificadas: preprocesamiento, selección de oraciones, selección de claves, formación de preguntas, generación de distractores y posprocesamiento.	Capacidad para manejar textos de conocimientos complejos Generación de distractores complejos, técnicas de evaluación estándar y datos de prueba estándar

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
12	Text Generation	Rich semantic representation-based approach for text generation	Ibrahim Fathy, Dalia Fadl and Mostafa Aref	2012	NLP, NLG, Text Summarization, Machine Translation and Information Retrieval applications	No se especifican técnicas de PLN	WordNet	En este artículo se propone un nuevo modelo para generar un texto en inglés a partir de RSG. El modelo propuesto se puede aprovechar en aplicaciones de resumen de texto, traducción automática y recuperación de información.	2	El modelo propuesto genera varios textos, utilizando ontología con WordNet Se describen las diferentes fases que son importantes en NLG	El modelo está orientado principalmente a generar ensayos del tipo narración y causa efecto No se describe la precisión de los resultados Los textos generados son oraciones cortas unidas para conformar un párrafo, llevando a pérdida de sentido.	-	Desarrollar un prototipo para realizar experimentos más complejos que el estudio de caso presentado Comparar resultados con modelos previamente relacionados
13	Question Answering	A computer-aided environment for generating multiple-choice test items	Ruslan Miktov, Le An Ha and Nikiforos Karamanis	2005	NLP, Automatic term extractionMultiple choice test itemsSentence transformationShallow parsing	Análisis sintáctico superficial, extracción automática de términos, transformación de oraciones, cálculo de la distancia semántica	WordNet	En este artículo se describe un nuevo procedimiento asistido por computadora para generar ítems de prueba de opción múltiple a partir de documentos electrónicos. Además de emplear varias técnicas de NLP	2	La metodología propuesta es general y se puede adaptar a diferentes campos o propósitos Se presenta una alternativa a la actividad de desarrollar manualmente preguntas de opción múltiple Se describe la manera en cómo realizar pruebas a través de experimentos con estudiantes aplicándoles los Items generados, además, uso de métricas con fórmulas	- El artículo cuenta con varios años (16 años), justificadamente implica que no se aprovechen los avances actuales de NLP y con ello obtener resultados de vanguardia	- Los resultados de la evaluación demuestra que la construcción de ítems de prueba de opción múltiple con la ayuda del programa desarrollado en el artículo es más eficiente que la producción puramente manual	- Incluir un extenso estudio basado en corpus para los géneros cubiertos cuyos hallazgos ayudaren a mejorar el esquema de clasificación existente descrito
14	Question Generation	Neural question generation from text: A preliminary study	Qingyu Zhou1, NanYang, Furu Wei, ChuanqiTan, Hangbo Bao and Ming Zhou	2018	Encoder-decoder, Heuristic rules, Natural languages, Neural Question Generation (NQG)	Etiquetas de parte del discurso (POS), Entidad con nombre (NER), Extracción de información, análisis de dependencias	SQuAD, GRU, BiGRU, BLEU	En este trabajo se propone aplicar el modelo codificador-decodificador neuronal para generar preguntas significativas y diversas a partir de oraciones en lenguaje natural, además, se lleva a cabo un estudio preliminar sobre la generación de preguntas a partir de texto con redes neuronales, denotado como NQG	3	Generación de preguntas fluidas y diversas a partir del texto	-	-	Investigar si las preguntas generadas automáticamente pueden ayudar a mejorar los sistemas de respuesta a preguntas
15	Text Generation	TextKD-GAN: Text Generation Using Knowledge Distillation and Generative Adversarial Networks	Md. Akmal Haidar and Mehdi Rezagholizadeh	2019	Text generation, Generative adversarial networks, Knowledge distillation	No se especifican técnicas de PLN	GAN, AE (Codificadores automáticos), SeqGAN, MaskGAN, LeakGAN, TextKD-GAN	En este trabajo se presenta un método que utiliza la destilación de conocimientos para explotar la configuración de GAN para la generación de texto. Se demuestra cómo se pueden utilizar los codificadores automáticos (AE) para proporcionar una representación continua de oraciones.	2	- Muestran el pseudocódigo del algoritmo TextKDGAN -	Los modelos basados en GAN están aún iniciando en tareas relacionadas con la generación de texto TextKD-GAN no es bueno generando oraciones largas	- Demuestran que TextKD-GAN supera a los generadores de texto convencionales basados en GAN que no necesitan capacitación previa.	Investigar el desempeño del enfoque en el nivel basado en palabras Mejorar TextKDGAN utilizando generadores basados en RNN
16	Automatic Generation Questions	Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams	Hill, Jennifer and Simha, Rahul	2016	Natural language processing, fill-in-theblank (FITB),	Word Cooccurrences	N-grams, cooccurrence likelihoods, WordNet	En este artículo se propone un método para generar automáticamente ejercicios de relleno de espacios en blanco de opción múltiple a partir de pasajes de texto existentes que desafían las habilidades de comprensión y la conciencia contextual del lector	4	Generación automáticamente de preguntas fill-in-the-blank (FITB) a partir de texto existente	Distractores inconsistentes semánticamente dado el texto completo Tener que filtrar los distractores adecuados	Éxito en generar preguntas con distractores que son semánticamente consistentes en un contexto estrecho	Explorar un método menos subjetivo para evaluar las palabras dentro de un contexto limitado Incorporar un proceso de filtrado de sinónimos más robusto, aprovechando los corpus ya utilizados Mejorar el rendimiento de generación de preguntas y demostrar la eficacia de este tipo de ejercicios para la mejora de habilidades de compresión lectora

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
17	Automatic Generation Questions	Automatic generation of English vocabulary tests	Yuni Susanti, Ryu Iida and Takenobu Tokunaga	2015	Automatic Question Generation, English Vocabulary Test, Word Sense Disambiguation, Multiple-choice Question, TOEFL Vocabulary Test	Word sense disambiguation (WSD)	WordNet, Context search (CS), Human Question (HQ), Machine Question (MQ)	En este artículo se presenta un modelo que, dada una palabra objetivo, genera un pasaje de lectura a partir de textos web recuperados de Internet, y luego emplear ese pasaje de lectura y el diccionario léxico de WordNet para generar opciones de preguntas (TOEFL)	4	- Especificación paso a paso de la estrategia utilizada para la generación de preguntas de opción múltiple, teniendo en cuenta la generación del pasaje, la respuesta correcta y los distractores	-	-	Mejorar las preguntas mal valoradas que genera el método a través de un análisis más detallado
18	Automatic Generation Questions	Evaluation of automatically generated English vocabulary questions	Yuni Susanti, Takenobu Tokunaga, Hitoshi Nishikawa and Hiroyuki Obari	2017	English vocabulary question, Automatic question generation, Multiple-choice question, Evaluation of question items, Neural test theory, Language learning Introduction	No se especifican técnicas de PLN	WordNet, Word sense disambiguation (WSD), Context search (CS), Human Question (HQ), Machine Question (MQ)	En este artículo se describen los detalles de los experimentos de evaluación para preguntas creadas por un sistema automático de generación de preguntas. Dada una palabra objetivo y uno de sus sentidos de palabra, el sistema genera una pregunta de vocabulario en inglés de opción múltiple que solicita el significado más cercano a la palabra objetivo en el pasaje de lectura.	3	Especificación a detalle del proceso realizado en el estudio, describiendo cada fase del estudio Descripción de métricas para la evaluación de preguntas generadas automáticamente por un sistema Datos y resultados relevantes del estudio	-	Demostración de la distinción entre preguntas generadas por máquinas y preguntas generadas por humanos, a través de métodos de evaluación sobre examinados.	Extender el sistema para generar otro tipo de preguntas y su evaluación Controlar la dificultad de las preguntas de vocabulario generadas automáticamente
19	Question Answering	Intelligent question answering system based on artificial neural network	Ahlam Ansari, Moonish Maknojia and Altamash Shaikh	2016	Question Answering System, NLP, Artificial Neural Network	Part of Speech (POS) tagging, Entity Recognition	Deep Neural Network (DNN)	En este artículo se propone un método para crear una red neuronal profunda a partir de los documentos proporcionados por usuarios y almacenarlos para uso futuro	2	Capacidad de responder preguntas complejas al asignar casos profundos a la palabra en las oraciones complicadas El sistema propuesto puede ser útil para una extracción de información más rápida por parte del usuario	-	-	- Extender el QAS a partir de extracción de imágenes, tablas y otros documentos estadísticos. Esto permitirá al usuario procesar más información y obtener la máxima extracción de QAS.
20	Question Answering	Question Answering Systems: Survey and Trends	Abdelghani BOUZHIANE, Djelloul BOUCHIHA, Noureddine DOUMI and Mimoun MALKI	2015	Question Answering System (QAS), Natural Language Processing (NLP), Information Retrieval, SPARQL, Semantic Web	Named entity recognition (NER)	-	En este artículo se clasifican diferentes QAS (Question Answering Systems) con el fin de tratar despejar el camino para investigadores en este ámbito, o en efecto, elegir el sistema adecuado a su problema. También pueden ver las deficiencias y corregirlas, o proponer nuevos sistemas de garantía de calidad.	2	Gráficas comparativas entre diferentes QAs Descripción de varios QAs con sus técnicas respectivas	Los sistemas mencionados son especialmente para respuestas exactas a preguntas concretas, basado en búsquedas en la Web No se formulan preguntas a partir de texto, sino que las preguntas son formuladas por el usuario, obteniendo una respuesta concreta	Logra guiar a los investigadores sobre los diferentes QAs existentes hasta esa fecha (2015), describiendo las ventajas de cada sistema y comparándolos con otros sistemas	Implementar un proyecto de QA para explorar datos vinculados. Formular la solicitud con el lenguaje natural árabe
21	Question Answering	Semantic question answering system using dbpedia	Passent M. ElKafrawy (&) , Amr M. Sauber, and Nada A. Sabry	2018	NLP, Question answering, Question classification , Dbpedia ontology, SVM, Sparql	Keyword extraction, Ngram (Unigram, Bigram), WHwords, Word shapes, POS tag, Related words, Chunk tags	AQUA, WordNet	En este artículo se presenta una arquitectura del sistema de respuesta a preguntas factoides utilizando la ontología Dbpedia. La arquitectura discutida se prueba y los resultados se comparan con los de otros sistemas.	3	- Usa Dbpedia que contiene información estructurada de Wikipedia, cuenta con millones de entidades, entre ellas resúmenes, coordenadas geográficas y representaciones, además el modelo de datos está representado en formato RDF (sujeto-predicado-objeto)	- Se desvía un poco del eje principal (Question Answering)	Se obtienen resultados aceptables en medida en que combinan una mayor cantidad de técnicas de NLP	- Encontrar un enorme conjunto de datos de preguntas y comparar los resultados con los obtenidos por otros sistemas. Luego usar este sistema para construir un sistema de respuesta a preguntas en árabe usando la ontología de Dbpedia árabe

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
22	Question Answering	Question Answering System on Education Acts Using NLP Techniques	Sweta P. Lende and Dr.M.M. Raghuwanshi	2015	Question Answering, NLP, Information Retrieval, Education Acts	Preprocessing: Stop words Removal, Stemming, POS tag Training: Tokenizing, stop word removal, Extracted keyword, stemming Index Term Dictionary	-	En este artículo se describen las diferentes metodologías y detalles de implementación del sistema de respuesta a preguntas para el lenguaje general y también se propone un Sistema de QA de dominio cerrado para el manejo de documentos relacionados con las secciones de actos educativos para recuperar respuestas más precisas utilizando técnicas de PNL	1	- Se describe con detalle la estructura del modelo propuesto en el artículo	- No presentan experimentos, ni resultados del modelo propuesto	-	No se especifica el trabajo futuro
23	Automatic Generation Questions	Evaluation of Automatically Generated Pronoun Reference Questions	Arief Yudha Satria and Takenobu Tokunaga	2018	Automatic question Generation, Natural Language Processing,	No se especifican técnicas de PLN	-	En este artículo se presenta la evaluación de preguntas de referencia de pronombres generadas automáticamente que preguntan a los examinados el antecedente del pronombre especificado en el pasaje de lectura	3	Descripción de cada parte del proceso Experimentos desde la perspectiva del profesor de inglés (evaluando si la pregunta puede ser incluida en un test, si la pregunta es aceptable con mejora, o si es aceptable sin ninguna mejora) Experimentos desde la perspectiva del estudiante de inglés (evaluando con tres conjuntos de preguntas que contienen diez preguntas generadas por máquinas y diez hechas por humanos)	No se mencionan técnicas de PLN ni información del modelo que usan en el sistema propuesto	-	Controlar la dificultad de los ítems de las preguntas generadas y generar otros tipos de preguntas
24	Text Generation	Semi-supervised neural text generation by joint learning of natural language generation and natural language understanding models	Raheel Qader, Franc Portet and Cyril Labbé	2019	Natural Language Processing (NLP), Natural Language Generation (NLG), Natural Language Understanding (NLU)	No se especifican técnicas de PLN	Seq2Seq, Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Recurrent Neural Networks (RNN)	En este documento, se propone un esquema de aprendizaje profundo semi-supervisado que puede aprender de datos no anotados y datos anotados cuando estén disponibles. Utiliza modelos de secuencia a secuencia de NLG y NLU	2	Resultados con tablas donde los datos fueron generados con métricas automáticas y evaluación humana Ejemplos que describen la entrada al algoritmo, la referencia, el modelo emparejado y modelo emparejado más modelo no emparejado	-	- El método utilizado logra resultados competitivos sin utilizar trucos de preprocesamiento o reclasificación	Reemplazar el ST Gumbel-Softmax con técnicas de aprendizaje por refuerzo como el gradiente de políticas Evaluar con otros mecanismos para aumentar el desempeño de no redundancia y cobertura de generación
25	Text Generation	Text Corpus for Natural Language Story-telling Sentence Generation: A Design and Evaluation	Worasa Limpanadusadee, Proadpran Punyabukkana, Atiwong Suchato. Onintra Poobrasert	2014	Natural Language Generation; N-Gram Model; Augmentative and Alternative Communication; Statistical Natural Language Processing; Learning Disabilities; Corpus Management	N-gram, POS Tagging	-	En este artículo, se informa sobre la creación de un corpus de texto de propósito especial destinado a ser utilizado para la generación natural de oraciones por AAC (Augmentative and Alternative Communication) en tareas de narración de historias por usuarios con discapacidades de aprendizaje	2	-	El texto generado es en idioma tailandés Problemas para obtener las oraciones objetivo correctas cuando el número de palabras era de 5 y 6 palabras	Precisión del modelo comparable con otros (BEST2010)	- Intentar incorporar más restricciones semánticas en el procedimiento de generación de oraciones para eliminar la ambigüedad de las posiciones probables de múltiples adjetivos y adverbios
26	Question Answering	Question Answering on Structured Data using NLIDB Approach	Vishal Wudaru., Nikhil Koditala, Aruneswara Reddy, Radhika Mamidi	2019	Natural Language Interface to Database (NLIDB), Natural Language Processing (NLP), Database Management System (DBMS), Query Language, Movie domain, Intermediate query approach	Tokenizer, Stop word remover, Dependency parsing, POS Tagging	-	En este artículo se presenta la construcción de un sistema de interfaz de lenguaje natural a base de datos (NLIDB) utilizando un enfoque de consulta intermedio. Este enfoque se demuestra utilizando el chatbot de dominio de película y también se puede extender a diferentes dominios.	2	El sistema se puede extender a otros dominios Simple y fácil de usar en comparación con el formulario entrada	Es un desafío estructurar la consulta en función del lenguaje natural del usuario El usuario no tiene conocimiento sobre qué tipo de preguntas el sistema podría abordar	El sistema puede abordar la mayoría de las consultas son películas cuya información existe en la base de datos. Se prueba manualmente y se determina que produce resultados precisos	- Extender el sistema para que funcione en varios idiomas, ampliar la base de datos y permitir que el sistema reconozca sinónimos de palabras usando un diccionario en línea como Word2Vec

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
27	Question Answering	Extracting Compact Sets of Features for Question Classification in Cognitive Systems: A Comparative Study	Marco Pota, Angela Fuggi, Massimo Esposito, Giuseppe De Pietro	2015	Cognitive Systems; NLP; Question Answering; Question Classification; Feature Extraction	Tokenization, POS Tagging, Stemming, Parsing	Support Vector Machine (SVM)	En este trabajo se proponen conjuntos compactos de características léxicas y sintácticas para distinguir las clases de preguntas. En particular, los unigramas ampliamente utilizados se sustituyen por un número menor de características, extraídas mediante la modificación de los procedimientos típicos de procesamiento del lenguaje natural para el análisis de preguntas	3	- Información amplia de la precisión del sistema (tablas, gráficos, otros), además, con diferentes combinaciones de características	-	Los resultados obtenidos por el algoritmo de aprendizaje de SVM en un conjunto de datos de referencia son coherentes con trabajos anteriores, y confirman que, dentro de las características léxicas y sintácticas, la combinación de unigramas madre y palabra principal permite obtener los mejores resultados en términos de precisión	Refinar la extracción de palabras de múltiples cabezas para reducir aún más su número
28	Question Answering	An Improved Automated Question Answering System from Lecture Videos	Divya Mamgai, Sonali Brodiya, Rohit Yadav and Mohit Dua	2019	Question answering system, NLP · Transcript, Language model	No se especifican técnicas de PLN	CMU Sphinx	En este documento se analiza la implementación de un sistema de garantía de calidad automatizado que utiliza una base de conocimientos construida a partir de las videoconferencias grabadas en las aulas, así como las que están disponibles en línea	2	-	Información poco detallada, descrita de manera superficial	Comparar diferentes CMU APIs y concluir cuál es la más adecuada para el sistema propuesto	Capacitar CMU Sphinx en otros idiomas y videos multilingües. El procesamiento de videos en términos de filtrado de ruido o mejora de la calidad del sonido se puede realizar antes del proceso de transcripción y también se puede automatizar para reducir el esfuerzo humano y mejorar la calidad del video.
29	Question Answering	Evaluating Reasoning in Factoid based Question Answering System by Using Machine Learning Approach	Ajitkumar Meshram Pundge and C. Namrata Mahender	2018	Question Answering, NLP, Information Retrieval, Term Frequency, Inverse Document Frequency, Cosine similarity	Tokenization, Stop Words, Term Frequency	Term frequency inverse document frequency, Cosine similarity; Scikitlearn	En este artículo se presenta un sistema de respuesta de preguntas de dominio cerrado centrado en preguntas de tipo por qué que pueden generar las respuestas de preguntas de tipo subjetivo	3	-	-	Evaluación del sistema dando la precisión, Recuperación y tasa de error del mismo, la precisión general del sistema presentado en el artículo es del 76,32%	- Trabajar en más preguntas de por qué y más lecciones agregadas y pruebas con más preguntas e intentaremos lograr la precisión que deseamos
30	Question Answering	A Recurrent BERTbased Model for Question Generation	Ying-Hong Chan and Yao-Chung Fan	2019	Natural Language Processing (NLP), Question Answering, Question Generation	Tokenizer, Word Embedding, Masked LM (MLM), Next Sentence Prediction (NSP)	Bidirectional Encoder Representations from Transformers (BERT), SQuAD, BLEU	En este estudio se investiga el empleo del modelo de lenguaje BERT previamente entrenado para abordar tareas de generación de preguntas. Presentan tres arquitecturas neuronales creadas sobre BERT para tareas de generación de preguntas.	5	El artículo describe detalladamente el modelo BERT, su estructura, su funcionamiento y presenta pruebas de experimento comparando los puntajes BLEU obtenidos con otros modelos El modelo alcanza un puntaje BLEU 4 que mejora el modelo anterior de BERT	-	Los resultados del experimento muestran que el modelo produce un rendimiento de vanguardia que avanza la puntuación BLEU 4 de los mejores modelos existentes de 16.85 a 22.17	No se especifica el trabajo futuro
31	Question Answering	An Approach for Extracting Exact Answers to Question Answering (QA) System for English Sentences	Raju Barskar, Gulfishan Firdose Ahmed, Nepal Barskar	2012	NLP, QA System; Predicate Logics, Knowledge Representation; Clause form, Herbrand's Theorem; Pattern Extracting	Named entity recognition (NER)	-	En este artículo se presenta un sistema de QA que aprovecha la información de la categoría explotando varios modelos de categorización de preguntas y respuestas	2	- Diagramas donde se especifica la estructura del sistema con la explicación de cada parte	- No se mencionan los experimentos y resultados sobre el sistema construido	-	No se especifica el trabajo futuro

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
32	Question Answering	Question categorization and classification using grammar-based approach	Alaa Mohasseb, Mohamed Bader-El-Den, Mihaela Cocea	2018	Question classification, Machine learning, Text mining, Text classification, Natural language processing (NLP)	Tokenization, POS tagging, Bag of Words	Decision Tree (DT), Naive Bayes, Support Vector Machine (SVM)	En esta investigación se presenta un análisis de los diferentes tipos de preguntas en función de su estructura gramatical, además, se identifican diferentes patrones y se usan algoritmos de aprendizaje automático para clasificarlos	3	Extensible a otro tipo de problemas de clasificación	-	Se logró proponer un enfoque formal basado en la gramática, los resultados muestran que la solución propuesta conduce a un buen desempeño en la clasificación de preguntas	Examinar y analizar más preguntas de diferentes conjuntos de datos y ampliar el análisis de los diferentes tipos de preguntas Investigar el uso del aprendizaje por conjuntos y otros métodos de minería de datos para tratar el problema de desequilibrio de clases
33	Question Answering	A Qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19	David Oniani and Yanshan Wang	2020	Covid-19, Cord-19, GPT-2, BERT, BioBERT, USE, TF-IDF, Dataset, Natural Language Processing (NLP), AI, Semantic Similarity	No se especifican técnicas de PLN	Term Frequency, Bidirectional Encoder Representations from Transformers (BERT), BioBERT (Bidirectional Encoder Representations from Transformers for Biomedical Text Mining), Universal Sentence Encoder (USE)	En este trabajo se propone aplicar un modelo de lenguaje para responder automáticamente preguntas relacionadas con COVID-19 y evaluar cualitativamente las respuestas generadas. Utilizan el modelo de lenguaje GPT-2 y se aplica el aprendizaje por transferencia para reentrenarlo en el corpus del conjunto de datos de investigación abierta COVID-19 (CORD-19)	3	- Comparación entre varios modelos que generan preguntas y respuestas, siendo BERT el modelo con mejor puntuación para la tarea en específico que presentan en el artículo	-	Los resultados fueron consistentes entre los anotadores, además, muestran que BERT logró el mejor desempeño al responder automáticamente las preguntas de COVID-19	No se especifica el trabajo futuro
34	Question Answering	Predicting Subjective Features from Questions on QA Websites using BERT	Issa Annamoradnejad, Mohammadamin Fazli, Jafar Habibi	2020	Q&A Websites; Subjective Features; Quality Prediction; Transfer Learning; BERT; NLP	No se especifican técnicas de PLN	Bidirectional Encoder Representations from Transformers (BERT), HuggingFace	En esta investigación se planea proporcionar un modelo que pueda predecir 20 aspectos subjetivos o de calidad de las preguntas utilizando BERT. Dado que estos aspectos incluyen preguntas sobre opiniones, recomendaciones o experiencias personales, son más difíciles de responder por computadora que las preguntas con respuestas únicas y verificables	2	-	-	El modelo logró Mean Squared Error (MSE) con un valor de 0.046 en la predicción de valores objetivo. Los resultados confirman que con un simple ajuste fino del modelo BERT previamente entrenado, podemos lograr una alta precisión, en poco tiempo y con menos cantidad de datos.	No se especifica el trabajo futuro
35	Text Generation	Text Generation for Imbalanced Text Classification	Suphamongkol Akkaradamrongrat, Pornpimon Kachamas, Sukree Sinthupinyo	2019	imbalanced text classification, text generation, Markov chains, LSTM	Tokenization, Word embedding	Markov Chains, Long Short-Term Memory (LSTM)	En esta investigación se utilizan técnicas de generación de texto para generar muestras sintéticas de clases minoritarias para equilibrar el conjunto de datos de texto. Se aplican dos métodos de generación de texto: la generación de texto usando cadenas de Markov y la generación de texto usando LSTM	2	Descripción completa del modelo, representado en diagramas y tablas	Los textos generados son poco naturales, tanto con Markov chains y Long Short Term Memory LSTM El texto generado es en tailandés	Se descubrió que Markov Chains superó tanto a la técnica de generación de texto LSTM como a la técnica de referencia, que es la técnica de sobre muestreo. Agregar textos de clase positivos generados por las cadenas de Markov podría hacer que el modelo de clasificador prediga una clase más positiva.	No se especifica el trabajo futuro
36	Text Generation	N-Gram Based Paraphrase Generator from Large text Document	Ashwini I Gadag, B M Sagar	2016	N-gram, candidate paraphrase, reference paraphrase, paraphrase generator	N-grams	Evaluation (BLEU, NIST, METEOR, ROUGE), R-precision	Este artículo explora diferentes tipos de métricas de evaluación de extracción de n-gramas, como métricas de evaluación de traducción automática y resumen de texto y métricas de precisión R.	1	-	El algoritmo genera texto en idioma canarés La precisión del modelo es regular	El algoritmo propuesto es capaz de generar un 46% de paráfrasis para cualquier tipo de entrada de corpus, la precisión se basa en el número de palabras comunes entre la paráfrasis candidata y de referencia.	No se especifica el trabajo futuro

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
37	Text Generation	Customizable text generation via conditional text generative adversarial network	Jinyin Chen, Yangyang Wu, Chengyu Jia, Haibin Zheng, Guohan Huang	2020	Text generation Variable length Emotion label Conditional text generative adversarial network	Feature Extraction, One-Hot encoded, POS Tagging,	a conditional text generative adversarial network (CTGAN), Long Short Term Memory (LSTM), convolutional neural network (CNN), SentiWordNet	En este artículo se presenta un sistema de generación de texto, llamado red de confrontación generativa de texto condicional personalizable, que es capaz de generar contenido de texto diverso de longitud variable con etiquetas de emoción personalizables	3	Diagramas detallados de la estructura del modelo, su funcionamiento y algoritmos en pseudocódigo Métricas de evaluación detalladas, tablas y gráficas con los resultados obtenidos	-	Se logró mostrar específicamente cómo CTGAN genera texto y cómo mejorar de manera efectiva su autenticidad, especialmente para reseñas de restaurantes y contenido de noticias. Finalmente, a través de varios indicadores de evaluación del texto generado, se compararon los resultados para resaltar la superioridad de CTGAN.	Aplicar CTGAN en aplicaciones de procesos de lenguaje más natural, como sistemas de diálogo y traducción automática.
38	Question Answering	Deep learning based question answering system for intelligent humanoid robot	Widodo Budiharto, Vincent Andreas and Alexander Agung Santoso Gunawan	2020	Humanoid Robot, NLP, Big data, Deep learning	Word Embedding, (GloVe)	Recurrent Neural Network (RNN) encoder, Convolution Neural Network (CNN) encoder, with Bidirectional Attention Flow (BiDAF), Google API Speech Recognition, SQuAD	En esta investigación se propone un Robot Humanoide Inteligente con capacidad de autoaprendizaje para aceptar y dar respuestas de personas basadas en la base de conocimientos de Deep Learning y Big Data	1	Se obtienen puntajes de evaluación aceptables en comparación con la red neuronal convolucional (CNN)	-	El uso de un codificador basado en RNN con BiDAF da una puntuación más alta que el codificador CNN con BiDAF. Se obtiene una puntuación F1 del 82,43% y el codificador basado en RNN dará una puntuación EM/F1 más alta que con el codificador CNN	Implementar una base de datos para guardar el conocimiento, de modo que el conocimiento pueda almacenar más datos y administrar fácilmente Mejorar el algoritmo para obtener mejores resultados en la respuesta a preguntas y manejar preguntas sin respuesta utilizando el conjunto de datos SQuAD 2.0
39	Question Generation	Transformer-based End-to-End Question Generation	Luis Enrico Lopez, Diane Kathryn Cruz, Jan Christian Blaise Cruz, Charibeth Cheng	2020	Natural Language Processing (NLP), Question Generation (QG)	No se especifican técnicas de PLN	Stanford Question Answering Dataset (SQuAD), GPT-2, BLEU, METEOR, ROUGE L	En este trabajo se muestra que las técnicas de ajuste fino basadas en transformadores se pueden utilizar para crear sistemas robustos de generación de preguntas utilizando un solo modelo de lenguaje previamente entrenado, sin el uso de mecanismos adicionales, metadatos de respuesta y funciones extensas.	5	- Rendimiento de vanguardia hasta el punto de superar modelos basados en seq2seq sin necesidad de características, técnicas y pasos de capacitación adicionales.	- Entrenado con el modelo más pequeño de GPT-2 debido a limitaciones de recursos computacionales	Se logra demostrar que el modelo propuesto funciona a la par con los modelos Seq2Seq que emplean la conciencia de respuesta y otros mecanismos especiales Se analizan los modos de falla del modelo y se identifica las posibles razones por las que falla el modelo	Entrenar modelos de generación de preguntas utilizando los tres tamaños de modelo GPT-2 más grandes
40	Text Generation	Texygen: A Benchmarking Platform for Text Generation Models	Yaoming Zhu, Sidi Lu, Lei Zheng, Jiaxian Guo, Weinan Zhang, Jun Wang, Yong Yu	2018	Text Generation, Benchmarking, Evaluation Metrics	No se especifican técnicas de PLN	Generative Adversarial Networks (GAN), Vanilla MLE, BLEU, EmbSim, NLL-Oracle, NLL-test, Self-BLEU	En este documento se presenta Texygen, una plataforma de evaluación comparativa de código abierto para modelos de generación de texto. Texygen incluye una variedad de métricas que evalúan la diversidad, calidad y consistencia de los textos generados.	4	Texygen proporciona un sistema de evaluación multidimensional estándar de arriba hacia abajo para modelos de generación de texto. Texygen proporciona el repositorio de código abierto de la plataforma, en el que los investigadores pueden encontrar la especificación y el manual de las API para implementar sus modelos para que Texygen los evalúe.	-	-	- Continuar actualizando nuevos modelos y diseñando métricas novedosas para una mejor evaluación comparativa de las tareas de generación de texto

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
41	Text Generation	Text Generation from Keywords	Kiyotaka Uchimoto, Satoshi Sekine, Hitoshi Isahara	2002	Text Generation, Keywords- Headwords, NLP	Bag of Words	-	En este artículo se propone un método para generar oraciones a partir de "keywords" o "headwords". Este método consta de dos partes principales, construcción y evaluación del texto candidato.	2	- El modelo puede servir para entrenarse en otros idiomas y así, generar texto en ese idioma	- El modelo está entrenado para el idioma Japonés	-	Generar una mayor variedad de oraciones de texto candidato Expandir las reglas que pueden generar un árbol de dependencia Preparar reglas predeterminadas o clasificar palabras en varias clases cuando no se encuentran oraciones que incluyan las palabras clave en el corpus de destino Ampliar la definición de palabras clave a otras palabras de contenido y sinónimos de las palabras clave
42	Text Generation	Language Generation with Recurrent Generative Adversarial Networks without Pre-training	Ofir Press, Amir Bar, Ben Bogin, Jonathan Berant, Lior Wolf	2017	Text Generation, Language Generation, Generative Adversarial Networks (GANs)	N-grams	Generative Adversarial Networks (GAN), Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU)	En este trabajo se muestra que se pueden entrenar redes neuronales recurrentes para generar texto con GAN desde cero utilizando el aprendizaje curricular, enseñando lentamente el modelo para generar secuencias de longitud creciente y variable.	2	Código abierto disponible en GitHub	-	-	- Aplicar estos modelos a tareas tales como subtítulos de imágenes y traducción, comparándolos con modelos entrenados con la máxima probabilidad
43	Text Generation	BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding	Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova	2019	Language Understanding, Transformers, Natural Language Processing (NLP)	WordPiece, Masked LM, Next Sentence Prediction (NSP)	Bidirectional Encoder Representations from Transformers (BERT), SQuAD	En este artículo se presenta un nuevo modelo de representación de lenguaje llamado BERT, que significa Representaciones de codificador bidireccional de Transformers, además se mejora los enfoques basados en el ajuste fino al proponer BERT	5	BERT es conceptualmente simple y empíricamente poderoso BERT obtiene nuevos resultados de vanguardia en once tareas de procesamiento del lenguaje natural Artículo sumamente detallado sobre BERT	-	Aumento de la puntuación GLUE al 80,5% (mejora absoluta del 7,7%), la precisión de MultiNLI al 86,7% (mejora absoluta del 4,6%), SQuAD v1.1 pregunta de respuesta Prueba F1 a 93,2 (mejora absoluta de 1,5 puntos) y SQuAD v2.0 Prueba F1 a 83,1 (mejora absoluta de 5,1 puntos).	No se especifica el trabajo futuro
44	Text Generation	Generating Text with Recurrent Neural Networks	Ilya Sutskever, James Martens, Geoffrey Hinton	2011	Text generation, Natural Language Processing (NLP), Artificial Neural Networks	N-grams	Recurrent Neural Network (RNN), Multiplicative Recurrent Neural Network (MRNN), Markov model	En este artículo se demuestra el poder de los RNN entrenados con el nuevo optimizador Hessian-Free (HF) aplicándolos a la tarea de predecir el siguiente carácter en un flujo de texto.	2	-	Puramente teórico	- Multiplicative Recurrent Neural Network (MRNN) predice el conjunto de prueba con mayor precisión que el memorizador de secuencias, pero con menos precisión que el PAQ sin diccionario en los tres conjuntos de datos.	No se especifica el trabajo futuro
45	Text Generation	Long Text Generation via Adversarial Training with Leaked Information	Jiaxian Guo, Sidi Lu, Han Cai, Weinan Zhang, Yong Yu, Jun Wang	2018	Text generation, Natural Language Processing (NLP), Leaked Information	No se especifican técnicas de PLN	Generative Adversarial Nets (GAN), LeakGAN, Long Short Term Memory (LSTM), Convolutional Neural Network (CNN). COCO, BLEU	En este artículo, se propone un nuevo marco, llamado LeakGAN, para abordar el problema de la generación de textos largos, permitiendo que la red discriminativa filtre sus propias características extraídas de alto nivel a la red generativa para ayudar aún más a la orientación.	3	LeakGAN es muy eficaz en la generación de texto largo y también mejora el rendimiento en escenarios de generación de texto corto. LeakGAN produce una mayor ganancia de rendimiento cuando se generan oraciones más largas.	-	LeakGAN logró una mejora significativa en el rendimiento con respecto a las soluciones anteriores, tanto en las puntuaciones BLEU como en las calificaciones humanas	Aplicar LeakGAN en aplicaciones de procesos de lenguaje más natural, como sistemas de diálogo y subtítulos de imágenes Mejorar la capacidad del discriminador para comprobar la coherencia global de toda la oración

Información Artículos - Generación de Texto y Respuesta a Preguntas Procesamiento del Lenguaje Natural													
Id	Tema	Nombre del documento	Autores	Año	Palabras clave	Técnicas PLN	Info. adicional (Red neuronal, Dataset, estrategias, etc)	Resumen	Relevancia (1-5)	Ventajas	Desventajas	Logros	Trabajos futuros
46	Text Generation	MASS: Masked Sequence to Sequence Pre-training for Language Generation	Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, TieYan Liu	2019	Text generation, Natural Language Processing (NLP)	No se especifican técnicas de PLN	Masked Sequence to Sequence pre-training (MASS), Bidirectional Encoder Representations from Transformers (BERT), GPT, neural Machine Translation (NMT),	En este artículo, inspirado en BERT, se propone un objetivo novedoso para la formación previa: el aprendizaje Masked Sequence to Sequence (MASS) para la generación de lenguajes.	3	En el artículo se incluye experimentos para diferentes tareas en las que el modelo puede aplicar	-	- MASS logró mejoras significativas sobre la línea de base sin entrenamiento previo o con otros métodos de entrenamiento previo. Más específicamente, MASS logró los puntajes BLEU de vanguardia para NMT sin supervisión	Aplicar MASS a más tareas de generación de lenguaje, como parafraseo de oraciones, transferencia de estilo de texto y edición posterior Investigar más del análisis teórico y empírico
47	Text Generation	Paraphrase Generation with Deep Reinforcement Learning	Zichao Li, Xin Jiang, Lifeng Shang, Hang Li	2018	Inverse problems, Learning systems, Natural language processing systems, Reinforcement learning	No se especifican técnicas de PLN	Seq2Seq, Attention Model, Teaching Forcing, ROUGE	En este artículo se presenta un enfoque de aprendizaje por refuerzo profundo para parafrasear la generación. Específicamente, se propone un nuevo marco para la tarea, que consiste en un generador y un evaluador, los cuales se aprenden de los datos.	3	Pseudocódigo de los algoritmos (RbM-SL y RbMIRL) Guía evaluación humana (Especificación de los criterios usados)	-	Los modelos propuestos (generadores) producen paráfrasis más precisas y superar a los métodos más avanzados en la generación de paráfrasis tanto en la evaluación automática como en la evaluación humana.	Aplicar el marco y las técnicas de capacitación en otras tareas, como la traducción automática y el diálogo.
48	Text Generation	Text Generation Service Model Based on Truth-Guided SeqGAN	Yuxi Wu, Junli Wang	2019	Text generation, generative adversarial networks, self-attention mechanism, truthguided	Word Segmentation, Feature Extraction	Generative Adversarial Networks (GAN), SeqGAN, Self-Attention, Convolution Neural Network (CNN), COCO	En este trabajo se estudia y analiza la tarea de generación de texto y se utiliza SeqGAN como modelo base	2	-	Puramente teórico	- Se mejoró la velocidad de convergencia del modelo de generación de texto con SeqGAN, y la calidad del texto generado por la red también se mejoró después de la estabilización, tanto en la pérdida de prueba NLL como en la similitud de incrustación	No se especifica el trabajo futuro
49	Text Generation	FGGAN: Feature-Guiding Generative Adversarial Networks for Text Generation	Yang Yang, Xiaodong Dan, Xuesong Qiu, Zhipeng Gao	2020	Generative Adversarial Networks, Text Generation, Deep Learning, Reinforcement Learning	No se especifican técnicas de PLN	Generative Adversarial Networks (GAN)	En este artículo se propone un marco mejorado FGGAN para la generación de texto, con el fin de resolver el problema de que la señal de retroalimentación del discriminador no es muy instructiva, este artículo propone un módulo de guía de características que obtiene la retroalimentación de características semánticas del texto al generador con más orientación.	3	-	Sí las reglas son complejas hará que el modo de la red neuronal colapse, es decir, algunas secuencias de texto estarán duplicadas Secuencias de texto cortas	FGGAN en este documento tiene algunas mejoras en el conjunto de datos relevante	Optimizar las reglas semánticas
50	Text Generation	Language Models are Unsupervised Multitask Learners	Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, OpenAI	2020	Neural Networks, Natural Language Processing, Unsupervised Multitask Learners, GPT	Lowercasing, Tokenization, Stemming, outof-vocabulary tokens	GPT-2, Transformers, Byte Pair Encoding (BPE), Conversation Question Answering dataset (CoQA)	En este artículo se demuestra que los modelos de lenguaje comienzan a aprender estas tareas sin ninguna supervisión explícita cuando se entrenan en un nuevo conjunto de datos de millones de páginas web llamado WebText	4	Resultados de vanguardia (GPT-2) en diferentes tareas Presentan al final del artículo experimentos (input, output) de diferentes tareas (Completion, conditional generation, summarization, translations, question answering)	El rendimiento de GPT-2 es mucho peor que el rango del 30 al 50% de los sistemas de respuesta a preguntas de dominio abierto que hibridan la recuperación de información con la respuesta a preguntas de documentos extractivos.	Realizar tareas posteriores en un entorno de cero posibilidades, sin ningún parámetro o modificación de la arquitectura. Se demuestra que este enfoque muestra potencial al resaltar la capacidad de los modelos de lenguaje para realizar una amplia gama de tareas en un entorno de cero oportunidades. Se logran resultados prometedores, competitivos y de vanguardia en función de la tarea.	Investigar el ajuste fino de puntos de referencia como decaNLP y GLUE

Referencias

1. Wang, Zheng & Wu, Qingbiao. (2018). An Integrated Deep Generative Model for Text Classification and Generation. *Mathematical Problems in Engineering*. 2018. 1-8. 10.1155/2018/7529286.
2. Jin, H., Cao, Y., Wang, T. et al. Recent advances of neural text generation: Core tasks, datasets, models and challenges. *Sci. China Technol. Sci.* 63, 1990–2010 (2020). <https://doi.org/10.1007/s11431-020-1622-y>
3. Park D., Ahn C.W. (2018) LSTM Encoder-Decoder with Adversarial Network for Text Generation from Keyword. In: Qiao J. et al. (eds) *Bio-inspired Computing: Theories and Applications. BIC-TA 2018. Communications in Computer and Information Science*, vol 952. Springer, Singapore. https://doi.org/10.1007/978-981-13-2829-9_35
4. Y. Qu, P. Liu, W. Song, L. Liu and M. Cheng, "A Text Generation and Prediction System: Pre-training on New Corpora Using BERT and GPT-2," 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 2020, pp. 323-326, doi: 10.1109/ICEIEC49280.2020.9152352.
5. S. Prabhumoye, P. Rai, L. S. Sandhu, Priya L and Sowmya Kamath S, "Automated query analysis techniques for semantics based question answering system," 2014 International Conference on Recent Trends in Information Technology, Chennai, 2014, pp. 1-6, doi: 10.1109/ICRTIT.2014.6996128.
6. Last, M., Danon, G. Automatic question generation (2020) *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10 (6), art. no. e1382, . 2-s2.0-85088593909
7. Dhall, I., Vashisth, S., Saraswat, S. Text Generation Using Long Short-Term Memory Networks (2020) *Lecture Notes in Networks and Systems*, 106, pp. 649-657
8. Kurtasov, Andrey. A System for Generating Cloze Test Items from {R}ussian-Language Text (2013) INCOMA Ltd. Shoumen, <https://www.aclweb.org/anthology/R13-2016>. Proceedings of the Student Research Workshop associated with {RANLP} 2013, pp. 107-112
9. Kurtasov, A. Automated generation of assessment test items from text: Some quality aspects (2014) *CEUR Workshop Proceedings*, 1197, pp. 91-95.

10. Jianglin, Yuan & Zhigang, Guo & Gang, Chen. (2019). Recurrent Convolution Attention Model (RCAM) for Text Generation based on Title. Journal of Physics: Conference Series. 1168. 052049. 10.1088/1742-6596/1168/5/052049.
11. Ch, D.R., Saha, S.K. Automatic Multiple Choice Question Generation from Text: A Survey (2020) IEEE Transactions on Learning Technologies, 13 (1), art. no. 8585151, pp. 14-25
12. Fathy, I., Fadl, D., Aref, M. Rich semantic representation based approach for text generation (2012) 2012 8th International Conference on Informatics and Systems, INFOS 2012, art. no. 6236603, pp. NLP20-NLP28.
13. Mitkov, R., Ha, L.A., Karamanis, N. A computer-aided environment for generating multiple-choice test items (2006) Natural Language Engineering, 12 (2), pp. 177-194
14. Zhou, Q., Yang, N., Wei, F., Tan, C., Bao, H., Zhou, M. Neural question generation from text: A preliminary study (2018) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10619 LNAI, pp. 662-671.
15. Haidar, M.A., Rezagholizadeh, M. TextKD-GAN: Text Generation Using Knowledge Distillation and Generative Adversarial Networks (2019) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11489 LNAI, pp. 107-118.
16. Hill, Jennifer & Simha, Rahul. (2016). Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams. 10.18653/v1/W16-0503.
17. Susanti, Y., Iida, R., Tokunaga, T. Automatic generation of english vocabulary tests (2015) CSEDU 2015 - 7th International Conference on Computer Supported Education, Proceedings, 1, pp. 77-87
18. Susanti, Y., Tokunaga, T., Nishikawa, H., Obari, H. Evaluation of automatically generated English vocabulary questions (2017) Research and Practice in Technology Enhanced Learning, 12 (1), art. no. 11
19. Ansari, A., Maknojia, M., Shaikh, A. Intelligent question answering system based on artificial neural network (2016) Proceedings of 2nd IEEE International Conference on Engineering and Technology, ICETECH 2016, art. no. 7569350, pp. 758-763
20. Bouziane, A., Bouchiha, D., Doumi, N., Malki, M. Question Answering Systems: Survey and Trends (2015) Procedia Computer Science, 73, pp. 366-375.
21. ElKafrawy, P.M., Sauber, A.M., Sabry, N.A. Semantic question answering system using dbpedia (2018) Lecture Notes in Computer Science (including subseries

Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10868 LNAI, pp. 821-832.

22. Lende, S.P., Raghuwanshi, M.M. nQuestion answering system on education acts using NLP techniques (2016) IEEE WCTFTR 2016 - Proceedings of 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare, art. no. 7583963
23. Satria, Arief & Tokunaga, Takenobu. (2017). Evaluation of Automatically Generated Pronoun Reference Questions. 76-85. 10.18653/v1/W17-5008.
24. Qader, R., Portet, F., Labbé, C. Semi-supervised neural text generation by joint learning of natural language generation and natural language understanding models (2019) INLG 2019 - 12th International Conference on Natural Language Generation, Proceedings of the Conference, pp. 552-562.
25. Limpanadusadee, W., Punyabukkana, P., Suchato, A., Poobrasert, O. Text corpus for natural language story-telling sentence generation: A design and evaluation (2014) 2014 11th Int. Joint Conf. on Computer Science and Software Engineering: "Human Factors in Computer Science and Software Engineering" - e-Science and High Performance Computing: eHPC, JCSSE 2014, art. no. 6841846, pp. 80-85
26. Wudaru, V., Koditala, N., Reddy, A., Mamidi, R. Question Answering on Structured Data using NLIDB Approach (2019) 2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019, art. no. 8728487, pp. 1-4.
27. Pota, M., Fuggi, A., Esposito, M., De Pietro, G. Extracting Compact Sets of Features for Question Classification in Cognitive Systems: A Comparative Study (2015) Proceedings - 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, 3PGCIC 2015, art. no. 7424626, pp. 551-556.
28. Mamgai, D., Brodiya, S., Yadav, R., Dua, M. An Improved Automated Question Answering System from Lecture Videos (2019) Lecture Notes in Networks and Systems, 46, pp. 653-659.
29. Pundge, A.M., Namrata Mahender, C. Evaluating Reasoning in Factoid based Question Answering System by Using Machine Learning Approach (2018) Proceedings of the 3rd International Conference on Communication and Electronics Systems, ICCES 2018, art. no. 8724085, pp. 821-825.
30. Chan, Ying-Hong & Fan, Yao-Chung. (2019). A Recurrent BERT-based Model for Question Generation. 154-162. 10.18653/v1/D19-5821.

31. Barskar, Raju & Ahmed, Gulfishan & Barskar, Nepal. (2012). An Approach for Extracting Exact Answers to Question Answering (QA) System for English Sentences. *Procedia Engineering*. 30. 1187-1194. 10.1016/j.proeng.2012.01.979.
32. Mohasseb, Alaa & Bader-El-Den, Mohamed & Haig, Ella. (2018). Question categorization and classification using grammar based approach. *Information Processing & Management*. 54. 10.1016/j.ipm.2018.05.001.
33. Oniani, David & Wang, Yanshan. (2020). A Qualitative Evaluation of Language Models on Automatic Question-Answering for COVID-19. *ArXiv*.
34. Annamoradnejad, Issa & Fazli, Mohammadamin & Habibi, Jafar. (2020). Predicting Subjective Features from Questions on QA Websites using BERT.
35. Akkaradamrongrat, S., Kachamas, P., Sinthupinyo, S. Text Generation for Imbalanced Text Classification (2019) *JCSSE 2019 - 16th International Joint Conference on Computer Science and Software Engineering: Knowledge Evolution Towards Singularity of Man-Machine Intelligence*, art. no. 8864181, pp. 181-186.
36. Gadag, A.I., Sagar, B.M. N-gram based paraphrase generator from large text document (2016) *2016 International Conference on Computation System and Information Technology for Sustainable Solutions, CSITSS 2016*, art. no. 7779447, pp. 91-94.
37. Chen, J., Wu, Y., Jia, C., Zheng, H., Huang, G. Customizable text generation via conditional text generative adversarial network (2020) *Neurocomputing*, 416, pp. 125-135.
38. Budiharto, W., Andreas, V., Gunawan, A.A.S. Deep learning-based question answering system for intelligent humanoid robot (2020) *Journal of Big Data*, 7 (1), art. no. 77, .
39. Lopez, Luis & Cruz, Diane & Cruz, Jan Christian Blaise & Cheng, Charibeth. (2020). Transformer-based End-to-End Question Generation.
40. Zhu, Y., Lu, S., Zheng, L., Guo, J., Zhang, W., Wang, J., Yu, Y. Taxygen: A benchmarking platform for text generation models (2018) *41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018*, pp. 1097-1100.
41. Uchimoto, Kiyotaka & Sekine, Satoshi & Isahara, Hitoshi. (2002). Text Generation from Keywords. 10.3115/1072228.1072292.

42. Press, Ofir & Bar, Amir & Bogin, Ben & Berant, Jonathan & Wolf, Lior. (2017). Language Generation with Recurrent Generative Adversarial Networks without Pre-training.
43. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding (2019) NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1, pp. 4171-4186.
44. Sutskever, I., Martens, J., Hinton, G. Generating text with recurrent neural networks (2011) Proceedings of the 28th International Conference on Machine Learning, ICML 2011, pp. 1017-1024.
45. Guo, J., Lu, S., Cai, H., Zhang, W., Yu, Y., Wang, J. Long text generation via adversarial training with leaked information (2018) 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, pp. 5141-5148.
46. Song, K., Tan, X., Qin, T., Lu, J., Liu, T.-Y. MASS: Masked sequence to sequence pre-training for language generation (2019) 36th International Conference on Machine Learning, ICML 2019, 2019-June, pp. 10384-10394
47. Li, Z., Jiang, X., Shang, L., Li, H. Paraphrase generation with deep reinforcement learning (2020) Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018, pp. 3865-3878.
48. Wu, Y., Wang, J. Text Generation Service Model Based on Truth-Guided SeqGAN (2020) IEEE Access, 8, art. no. 8957565, pp. 11880-11886.
49. Yang, Y., Dan, X., Qiu, X., Gao, Z. FGGAN: Feature-Guiding Generative Adversarial Networks for Text Generation (2020) IEEE Access, 8, art. no. 9091179, pp. 105217-105225.
50. A. Radford and Jeffrey Wu and R. Child and David Luan and Dario Amodei and Ilya Sutskever, Language Models are Unsupervised Multitask Learners (2019)