Advanced Topics in Machine Learning

Winter Semester 2024/2025

Prof. Dr.-Ing. Christian Bergler | OTH Amberg-Weiden

Topics of Today: Advanced Deep Learning Strategies – Part III

- Self-Supervised Learning
- Active Learning

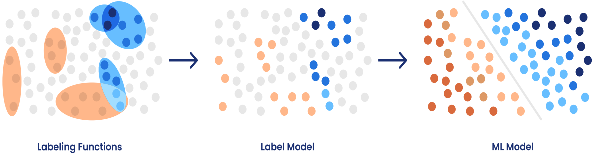
Topics of Today: Advanced Deep Learning Strategies – Part III

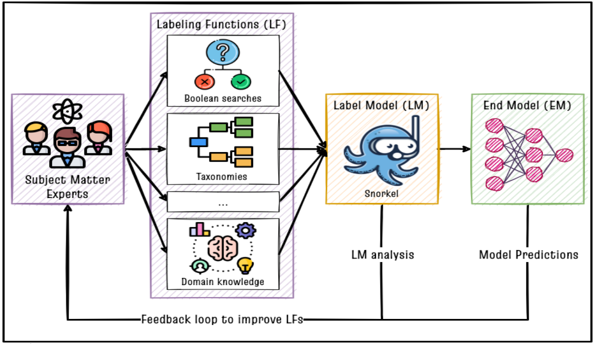
- Weak Supervision
- N-Shot Learning
- Causal Learning

Deep Learning Strategies – Part III

Weak Supervision

- Goal: How to transfer an unsupervised to a (weakly) supervised task!



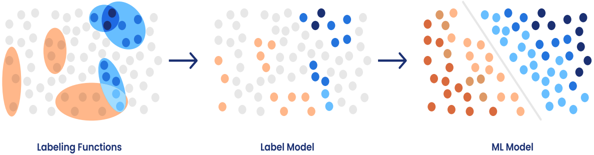


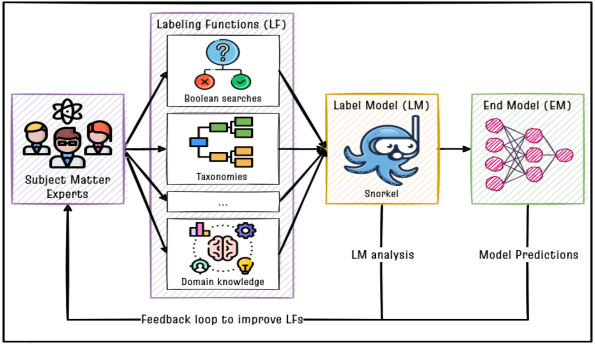
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Goal: How to transfer an unsupervised to a (weakly) supervised task!
- Interpret & categorize high-level, scalable, and potentially noisy signal sources by applying multiple sources of supervision





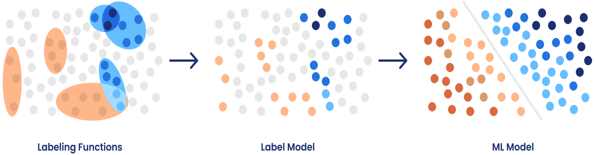
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

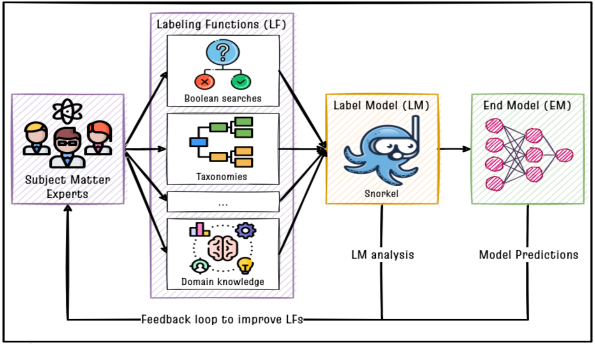
Deep Learning Strategies – Part III

OTH Amberg-Weiden

Weak Supervision

- Goal: How to transfer an unsupervised to a (weakly) supervised task!
- Interpret & categorize high-level, scalable, and potentially noisy signal sources by applying multiple sources of supervision
- Create large training sets, based on unlabeled data, much more quickly than using a 1-by-1 manual supervision



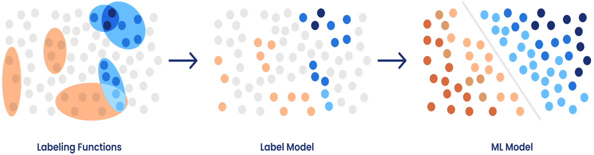


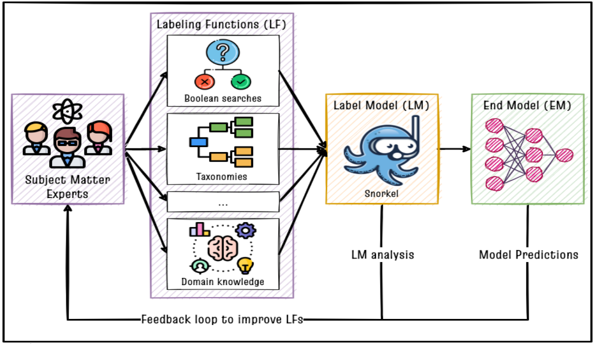
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Goal: How to transfer an unsupervised to a (weakly) supervised task!
- Interpret & categorize high-level, scalable, and potentially noisy signal sources by applying multiple sources of supervision
- Create large training sets, based on unlabeled data, much more quickly than using a 1-by-1 manual supervision
- Weak supervision training on incomplete, partially (machine-)labeled data





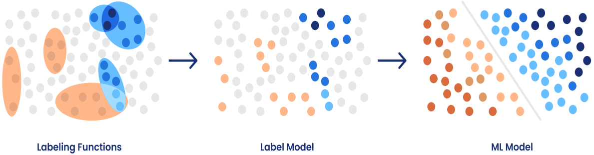
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

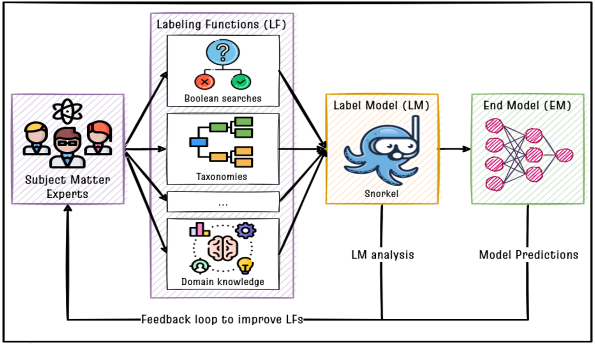
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Goal: How to transfer an unsupervised to a (weakly) supervised task!
- Interpret & categorize high-level, scalable, and potentially noisy signal sources by applying multiple sources of supervision
- Create large training sets, based on unlabeled data, much more quickly than using a 1-by-1 manual supervision
- Weak supervision training on incomplete, partially (machine-)labeled data
- Lack of precision: completely labeled, but often not precisely enough to recognize subtle differences





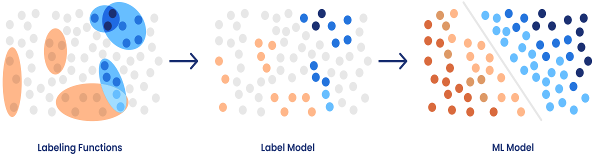
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

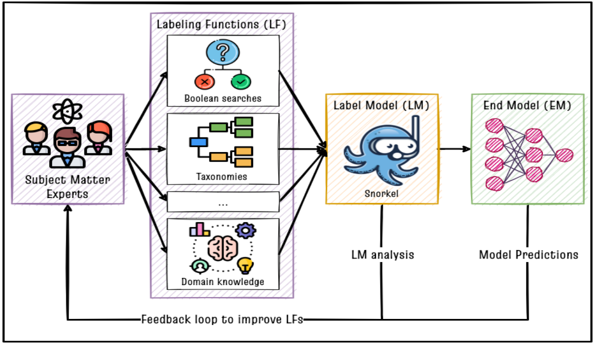
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Lack of accuracy: usually completely labeled, but some labels are wrong, in order to train robust ML-models





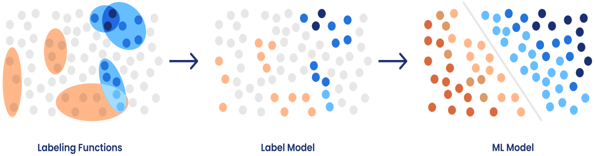
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

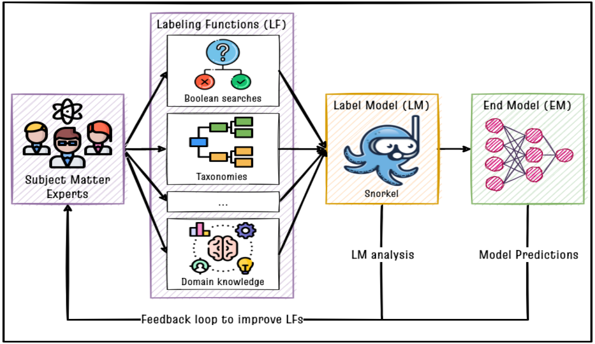
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Lack of accuracy: usually completely labeled, but some labels are wrong, in order to train robust ML-models
- Weakly supervised data is generated through crowdsourcing, heuristics, noisy predictions from pre-trained models, ...



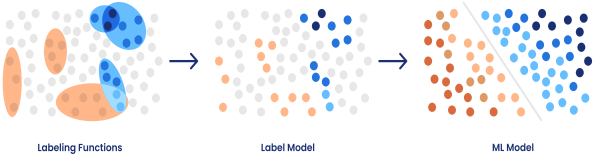


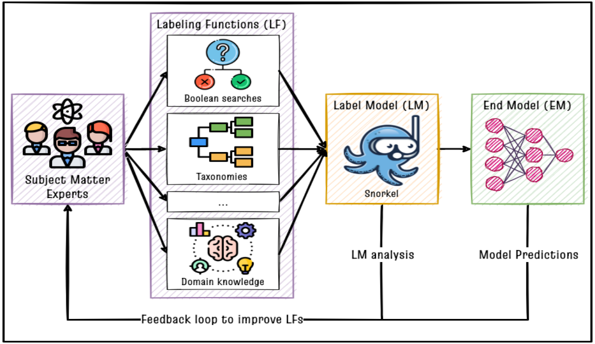
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Lack of accuracy: usually completely labeled, but some labels are wrong, in order to train robust ML-models
- Weakly supervised data is generated through crowdsourcing, heuristics, noisy predictions from pre-trained models, ...
- Different labeling functions λ (majority vote) $\rightarrow$ learn when, where, & how much to trust, learning their expertise and how to combine their votes $\rightarrow$ Highest possible label quality for each data point



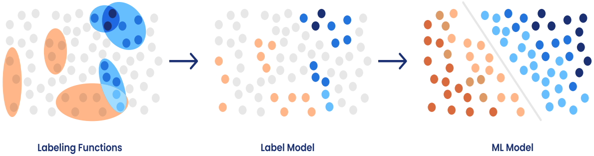


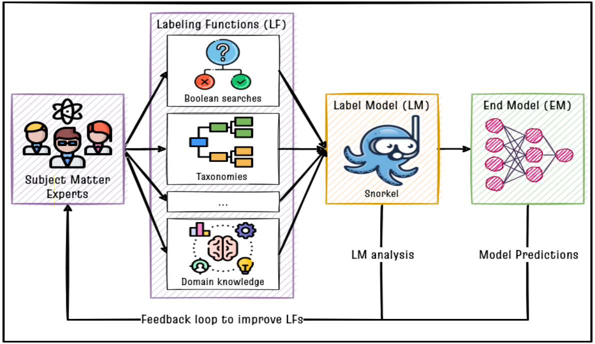
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Deep Learning Strategies – Part III

Weak Supervision

- Lack of accuracy: usually completely labeled, but some labels are wrong, in order to train robust ML-models
- Weakly supervised data is generated through crowdsourcing, heuristics, noisy predictions from pre-trained models, ...
- Different labeling functions λ (majority vote) $\rightarrow$ learn when, where, & how much to trust, learning their expertise and how to combine their votes $\rightarrow$ Highest possible label quality for each data point
- Idea: trade-off 100,000 pretty good vs. 100 perfect data labels





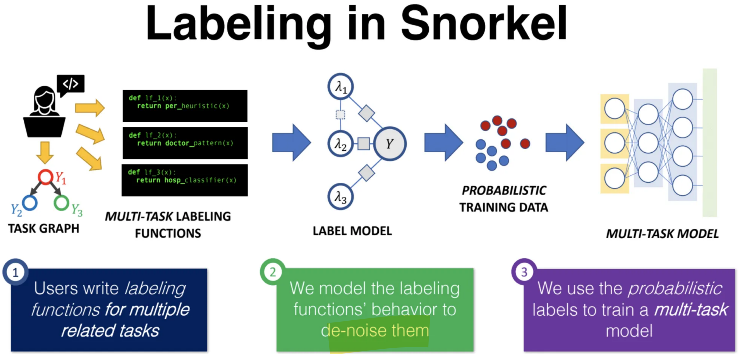
Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

Source: https://pub.towardsai.net/no-training-data-no-problem-weak-supervision-to-the-rescue-bd38b475a412

prob of diversity of 100 perfect is less than prob diversity of 100000 get rid of false classification: de noise

OTH Amberg-Weiden

Snorkel Framework



- Labeling functions λ_N (simple functions/rules), providing multiple inconsistent or overlapping labels (commercialized version available here: Snorkel)
- Combine noisy weak labels in a labeling model to generate a consensusmore accurate final label
- Label model learns dependencies/reliabilities of the labeling functions, weighting each function according to its accuracy and overlap with others
- Probabilistic labels for each sample, allow to build a (weakly) supervised training set

Source: Link to Image

OTH Amberg-Weiden

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

OTH Amberg-Weiden

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)
- Especially useful for weak supervision, where perfect labels are rare or costly to obtain

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)
- Especially useful for weak supervision, where perfect labels are rare or costly to obtain
- Heuristic-based Labeling: using domain-specific rules or heuristics to generate approximate labels (e.g. pattern matching, thresholds, ...)

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)
- Especially useful for weak supervision, where perfect labels are rare or costly to obtain
- Heuristic-based Labeling: using domain-specific rules or heuristics to generate approximate labels (e.g. pattern matching, thresholds, ...)
- Model-based Labeling: applying pre-trained models to derive labels for new data (e.g. large task-related ML/DL-models)

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Deep Learning Strategies – Part III
Weak Supervision

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)
- Especially useful for weak supervision, where perfect labels are rare or costly to obtain
- Heuristic-based Labeling: using domain-specific rules or heuristics to generate approximate labels (e.g. pattern matching, thresholds, ...)
- Model-based Labeling: applying pre-trained models to derive labels for new data (e.g. large task-related ML/DL-models)
- Clustering and Similarity Methods: unsupervised methods to group similar data points, which are then labeled belonging to specific class (k-means, cosine-similarity, ...)

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Deep Learning Strategies – Part III

Weak Supervision

Smart Labeling

- Labeling involves using automated or semi-automated techniques to label data intelligently (see pseudo-labeling, self-supervised approaches)
- Especially useful for weak supervision, where perfect labels are rare or costly to obtain
- Heuristic-based Labeling: using domain-specific rules or heuristics to generate approximate labels (e.g. pattern matching, thresholds, ...)
- Model-based Labeling: applying pre-trained models to derive labels for new data (e.g. large task-related ML/DL-models)
- Clustering and Similarity Methods: unsupervised methods to group similar data points, which are then labeled belonging to specific class (k-means, cosine-similarity, ...)
- Active Learning: label only the most uncertain or informative data points, making the labeling process more efficient

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will “cut in” or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will "cut in" or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)
- Cut-in prediction using weak supervision, because collection/labeling large amounts of data for every possible "cut-in" situation is challenging

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

OTH Amberg-Weiden

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will "cut in" or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)
- Cut-in prediction using weak supervision, because collection/labeling large amounts of data for every possible "cut-in" situation is challenging
- Models learn to recognize "noisy cut-in" maneuvers based on weak/noisy signals, rather than explicitly labeled data

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will "cut in" or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)
- Cut-in prediction using weak supervision, because collection/labeling large amounts of data for every possible "cut-in" situation is challenging
- Models learn to recognize "noisy cut-in" maneuvers based on weak/noisy signals, rather than explicitly labeled data
- Heuristics: patterns like relative vehicle speeds, distances, and lateral movements can be used to infer likely cut-in scenarios.

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will "cut in" or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)
- Cut-in prediction using weak supervision, because collection/labeling large amounts of data for every possible "cut-in" situation is challenging
- Models learn to recognize "noisy cut-in" maneuvers based on weak/noisy signals, rather than explicitly labeled data
- Heuristics: patterns like relative vehicle speeds, distances, and lateral movements can be used to infer likely cut-in scenarios.
- Sensor fusion: combining data from multiple sensors (e.g., LIDAR, radar, cameras) to create an accurate model that can predict "cut-in" even when labeling data is limited

Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

every potential event has to be taken into account

here patrren recognition is important

Application: Cut-In Prediction Autonomous Vehicle

- Refers to predicting when another vehicle in adjacent lanes will “cut in” or merge into the lane occupied by the autonomous vehicle (critical for safe driving!)
- Cut-in prediction using weak supervision, because collection/labeling large amounts of data for every possible “cut-in” situation is challenging
- Models learn to recognize “noisy cut-in” maneuvers based on weak/noisy signals, rather than explicitly labeled data
- Heuristics: patterns like relative vehicle speeds, distances, and lateral movements can be used to infer likely cut-in scenarios.
- Sensor fusion: combining data from multiple sensors (e.g., LIDAR, radar, cameras) to create an accurate model that can predict “cut-in” even when labeling data is limited
- Unlabeled data: process large amounts of unlabeled driving data with weakly supervised trained models that learn to identify “cut-in” scenarios indirectly

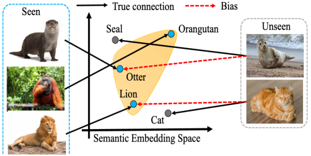
Source: Image from OTH-AW, Electrical Engineering, Media and Computer Science, Thomas Nierhoff – Vorlesung Advanced Topics in ML

train model on weakly supervissed data sets. from sensor data, heuristics etc and unlabeled

N-Shot Learning – Zero-Shot, Single-Shot & Few-Shot Learning

Idea: Zero-shot, single/one-shot, and few-shot learning are learning paradigms where models identify/categorize or generate new categories with very few or even no labeled examples

- Zero-Shot Learning: identify new categories never seen in training by leveraging auxiliary semantic knowledge based on feature similarity via textual/visual embeddings!



(a) a tapir made of accordion. a tapir with the texture of an accordion.

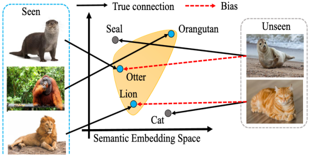
(b) an illustration of a baby hedgehog in a christmas sweater walking a dog

Source: https://medium.com/@vireshj/zero-shot-learning-understanding-machines-that-learn-like-humans-e670f83186b8
Source: Aditya Ramesh et al., "Zero-Shot Text-to-Image Generation", Figure 2

N-Shot Learning – Zero-Shot, Single-Shot & Few-Shot Learning

Idea: Zero-shot, single/one-shot, and few-shot learning are learning paradigms where models identify/categorize or generate new categories with very few or even no labeled examples

- Zero-Shot Learning: identify new categories never seen in training by leveraging auxiliary semantic knowledge based on feature similarity via textual/visual embeddings!
- Non-static set of classes, while unseen events will be stored in the high-dimensional feature space either near (related to "Knowns") or far-apart (unrelated to "Knowns")



(a) a tapir made of accordion. a tapir with the texture of an accordion.

(b) an illustration of a baby hedgehog in a christmas sweater walking a dog

Source: https://medium.com/@vireshj/zero-shot-learning-understanding-machines-that-learn-like-humans-e670f83186b8

Source: Aditya Ramesh et al., "Zero-Shot Text-to-Image Generation", Figure 2

Deep Learning Strategies – Part III

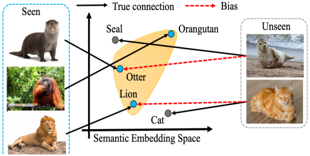
N-Shot Learning

zero shot for first time seeing few shot is for few times seeing

N-Shot Learning – Zero-Shot, Single-Shot & Few-Shot Learning

Idea: Zero-shot, single/one-shot, and few-shot learning are learning paradigms where models identify/categorize or generate new categories with very few or even no labeled examples

- Zero-Shot Learning: identify new categories never seen in training by leveraging auxiliary semantic knowledge based on feature similarity via textual/visual embeddings!
- Non-static set of classes, while unseen events will be stored in the high-dimensional feature space either near (related to “Knowns”) or far-apart (unrelated to “Knowns”)
- Class assignment for a first-time appearance of a new category requires pre-existing knowledge (embeddings, e.g. single-shot) describing the new class



(a) a tapir made of accordion. a tapir with the texture of an accordion.

(b) an illustration of a baby hedgehog in a christmas sweater walking a dog

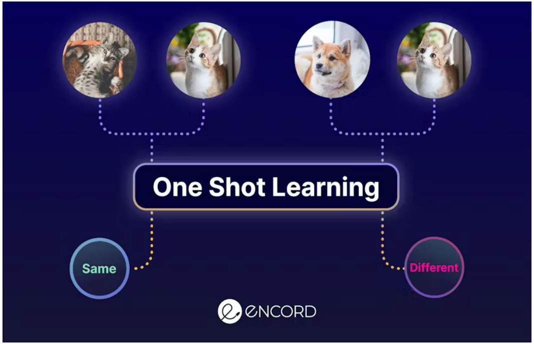
Source: https://medium.com/@vireshj/zero-shot-learning-understanding-machines-that-learn-like-humans-e670f83186b8

Source: Aditya Ramesh et al., “Zero-Shot Text-to-Image Generation”, Figure 2

Deep Learning Strategies – Part III

N-Shot Learning

- Single(=One)-Shot Learning: involves learning a new category from a single instance in training

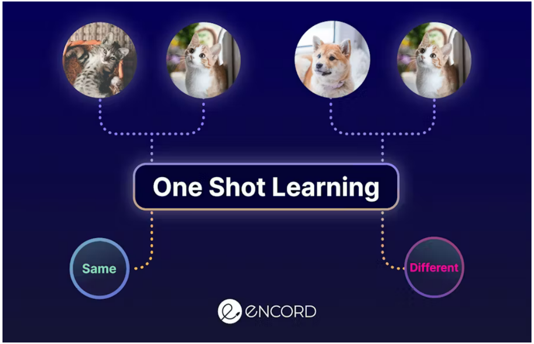


Source: https://encord.com/blog/one-shot-learning-guide/

Deep Learning Strategies – Part III

N-Shot Learning

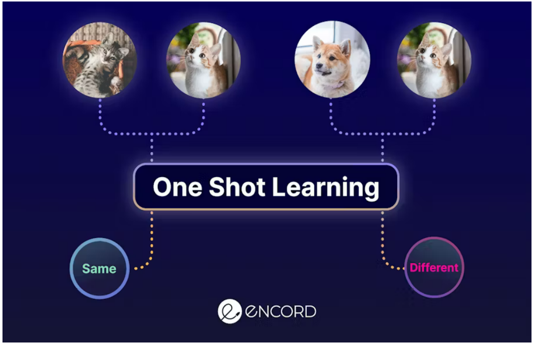
- Single(=One)-Shot Learning: involves learning a new category from a single instance in training
- Comparison of new data points w.r.t. the given single sample using similarity measures or embeddings in a latent space (deep metric learning!)



Source: https://encord.com/blog/one-shot-learning-guide/

Deep Learning Strategies – Part III
N-Shot Learning

- Single(=One)-Shot Learning: involves learning a new category from a single instance in training
- Comparison of new data points w.r.t. the given single sample using similarity measures or embeddings in a latent space (deep metric learning!)
- These networks learn to measure similarity between data samples, enabling classification with just a single instance

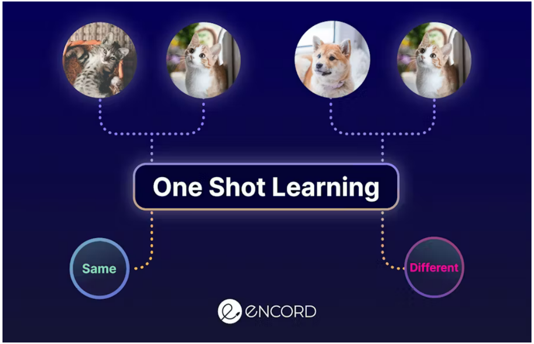


Source: https://encord.com/blog/one-shot-learning-guide/

Deep Learning Strategies – Part III

N-Shot Learning

- Single(=One)-Shot Learning: involves learning a new category from a single instance in training
- Comparison of new data points w.r.t. the given single sample using similarity measures or embeddings in a latent space (deep metric learning!)
- These networks learn to measure similarity between data samples, enabling classification with just a single instance
- Class assignment is based on distance-related outcome ("close enough") of an unseen sample w.r.t. the "single shot data point"

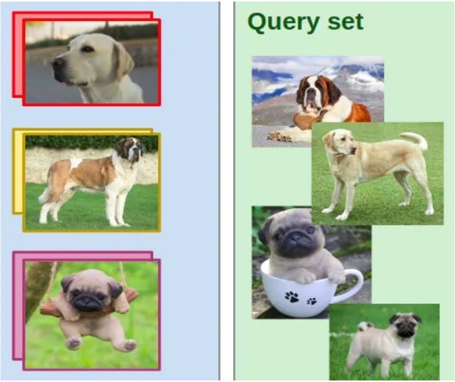


Source: https://encord.com/blog/one-shot-learning-guide/

Deep Learning Strategies – Part III

N-Shot Learning

- Few-Shot Learning: model learns from a small number of labeled examples, with more than zero/one images to compare against

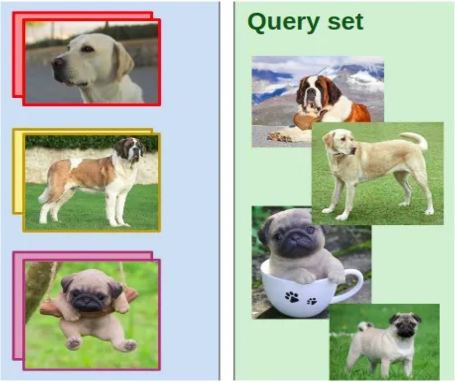


Source: https://neptune.ai/blog/understanding-few-shot-learning-in-computer-vision

Deep Learning Strategies – Part III

N-Shot Learning

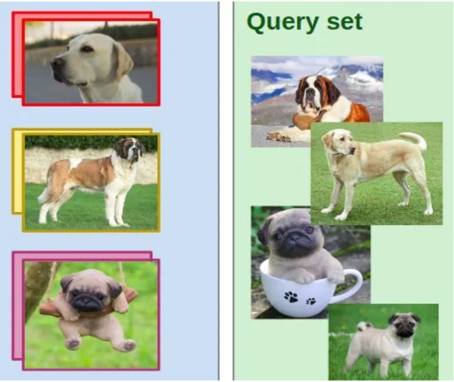
- Few-Shot Learning: model learns from a small number of labeled examples, with more than zero/one images to compare against
- Number of labeled "shots" $N < 10$, usually between 2 to 5, while model generalization is based on this small set of instances



Source: https://neptune.ai/blog/understanding-few-shot-learning-in-computer-vision

OTH Amberg-Weiden

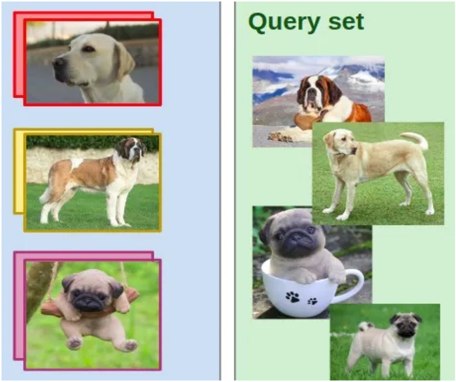
- Few-Shot Learning: model learns from a small number of labeled examples, with more than zero/one images to compare against
- Number of labeled "shots" $N < 10$, usually between 2 to 5, while model generalization is based on this small set of instances
- Often a K-way N-Shot classification problem, with K as the number of classes and N as the class-specific small number of labels, together with Q many query images



Source: https://neptune.ai/blog/understanding-few-shot-learning-in-computer-vision

causal learning home work

- Few-Shot Learning: model learns from a small number of labeled examples, with more than zero/one images to compare against
- Number of labeled "shots" $N < 10$, usually between 2 to 5, while model generalization is based on this small set of instances
- Often a K-way N-Shot classification problem, with K as the number of classes and N as the class-specific small number of labels, together with Q many query images
- Classify Q query images among the K classes, with $K \times N$ samples in the training set are the only representatives

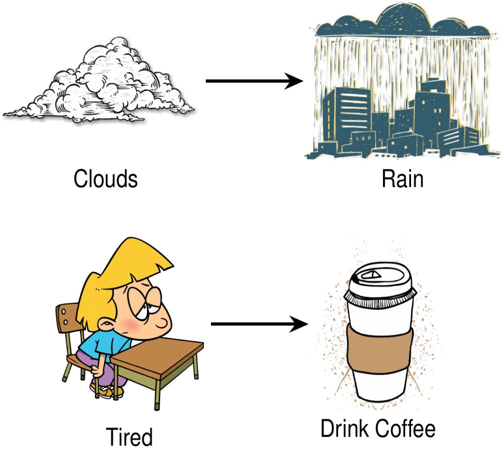


Source: https://neptune.ai/blog/understanding-few-shot-learning-in-computer-vision

Deep Learning Strategies – Part III

Causal Learning

- Implies a direct cause-and-effect relationship ("explain the cause of an effect") between two variables (change one variable will directly impact the other in a predictable way)



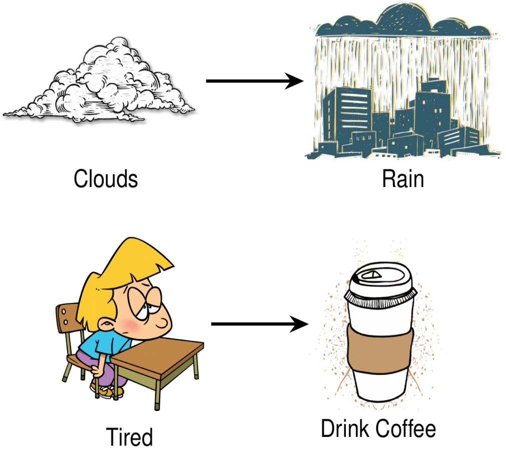
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

OTH Amberg-Weiden

- Implies a direct cause-and-effect relationship (“explain the cause of an effect”) between two variables (change one variable will directly impact the other in a predictable way)
- Careful: correlation describes a statistical association between two variables and indicates a simultaneous change (trend!) $\rightarrow$ However, not necessarily causing the other!

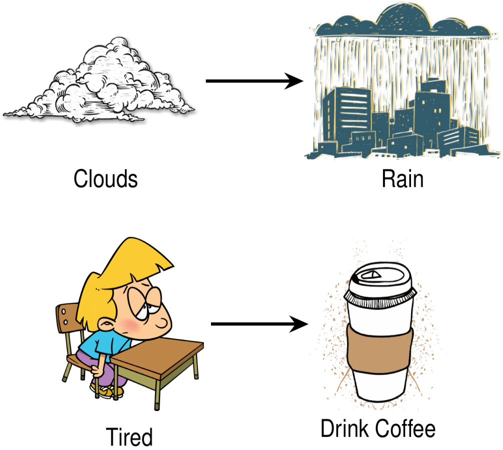


Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

- Implies a direct cause-and-effect relationship ("explain the cause of an effect") between two variables (change one variable will directly impact the other in a predictable way)
- Careful: correlation describes a statistical association between two variables and indicates a simultaneous change (trend!) $\rightarrow$ However, not necessarily causing the other!
- Causal inference often uses statistical methods to identify relationships that are not simply correlations but have a directional influence



Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Types of Causal Relationships

- Direct Causal Relationships: One variable directly affects another (e.g., smoking directly causes an increased risk of lung cancer)
- Indirect Causal Relationships: The effect of one variable on another is mediated through a third variable (e.g., education may lead to higher income, and higher income might then lead to better health outcomes)
- Bidirectional Causality: Some relationships may involve mutual causality, where two variables affect each other (e.g., diet and exercise can both affect each other over time)

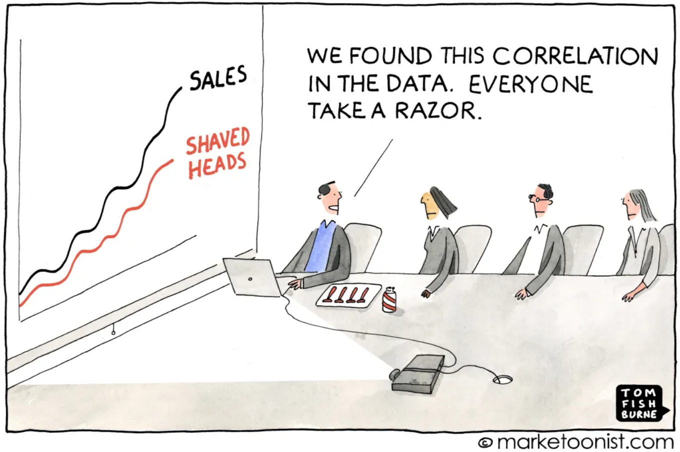
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Causal Learning

Source: Image from Nan Rosemary Ke, Stefan Bauer, "Causality and Deep Learning: Synergies, Challenges and the Future", Slide 29

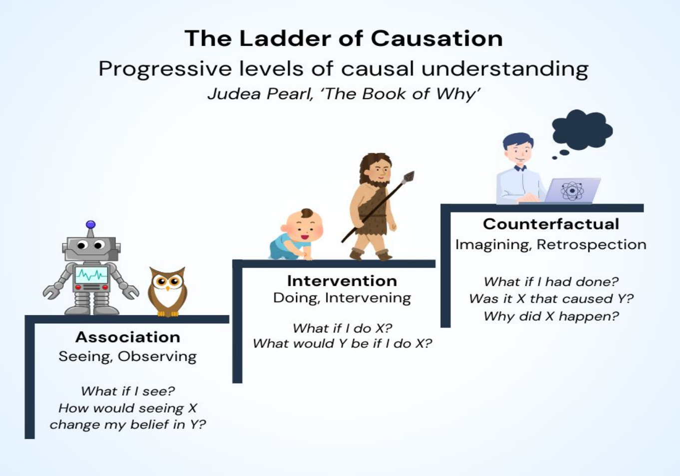
OTH Amberg-Weiden

Correlation is NOT Causation!



Source: Image from https://censemaking.com/2022/08/03/the-fear-and-folly-of-data/

Ladder of Causation



- Organizes causal reasons into three hierarchical levels: association, intervention, counterfactuals
- From simply observing data to fully understanding & reasoning about causal mechanisms

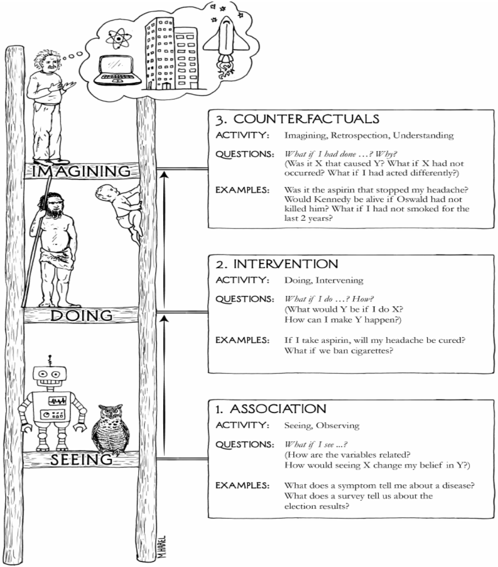
Source: Image from https://www.linkedin.com/posts/causai-bv_ladderofcausation-causalai-decisionmaking-activity-7177992092458057729-hiP9

Deep Learning Strategies – Part III

Causal Learning

Association

- Observing correlations/dependencies of variables: "What happens if we change this variable?"



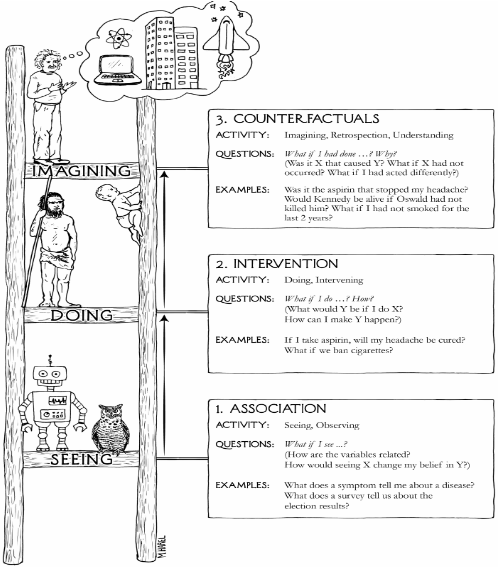
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Association

- Observing correlations/dependencies of variables: “What happens if we change this variable?”
- However, no explanation why one thing happens because of another (purely observational!)



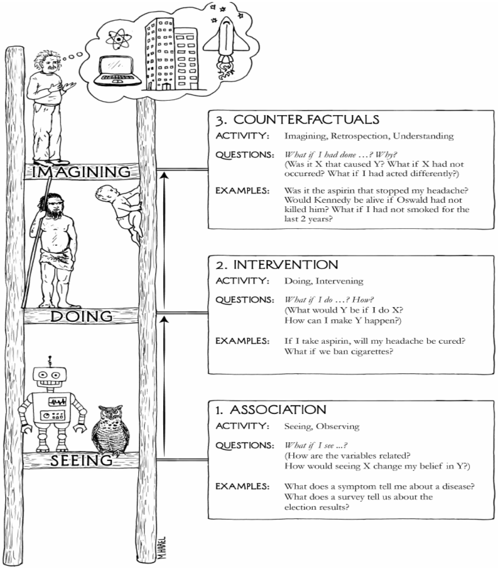
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Association

- Observing correlations/dependencies of variables: "What happens if we change this variable?"
- However, no explanation why one thing happens because of another (purely observational!)
- Questions:"What is the probability of Y given X?



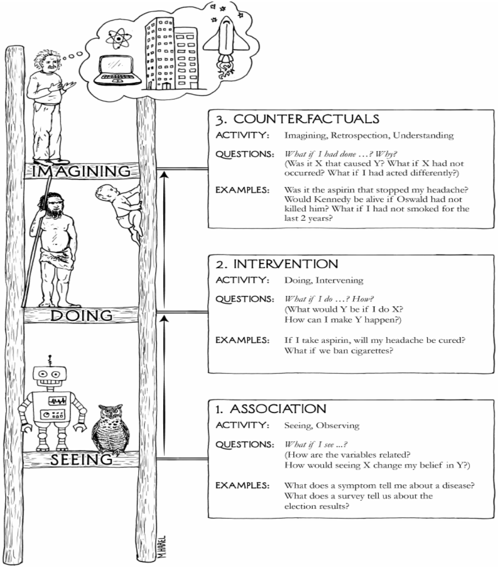
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Association

- Observing correlations/dependencies of variables: “What happens if we change this variable?”
- However, no explanation why one thing happens because of another (purely observational!)
- Questions:“What is the probability of Y given X?
- Methods: traditional statistical methods, ML (linar regression, classification, clustering)

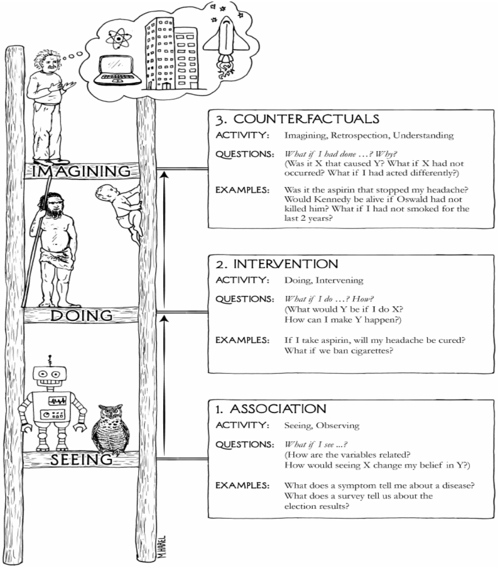


Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III
Causal Learning

Association

- Observing correlations/dependencies of variables: "What happens if we change this variable?"
- However, no explanation why one thing happens because of another (purely observational!)
- Questions:"What is the probability of Y given X?
- Methods: traditional statistical methods, ML (linar regression, classification, clustering)
- Association alone do not determine if a relationship is causal (correlation between ice cream sales and drowning incidents does not imply that buying ice cream causes drowning)



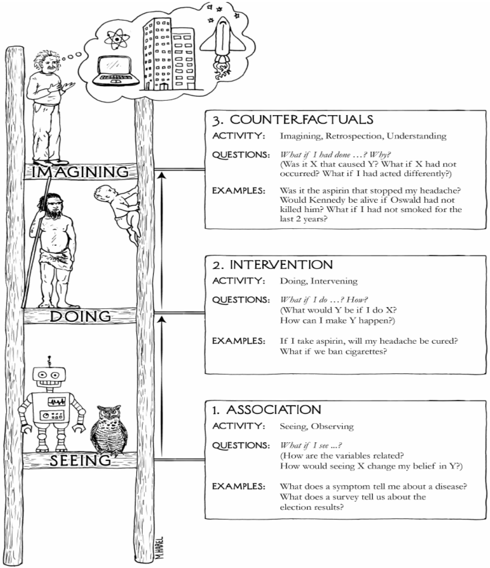
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Intervention

- Moving beyond observation by introducing actions to change the system and observing the results: "What if I do this?"



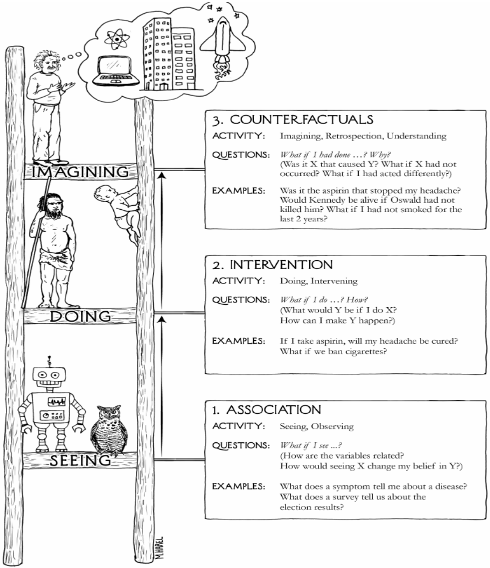
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Intervention

- Moving beyond observation by introducing actions to change the system and observing the results: "What if I do this?"
- Represent causal relationship by showing how changes in one variable affect another



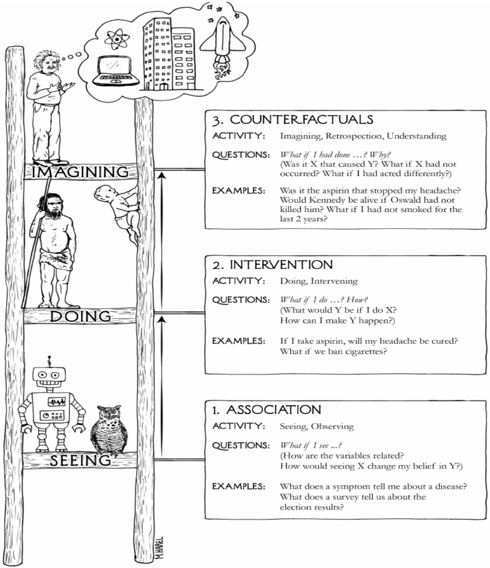
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Intervention

- Moving beyond observation by introducing actions to change the system and observing the results: "What if I do this?"
- Represent causal relationship by showing how changes in one variable affect another
- Question: "What will happen to Y if I change X?"



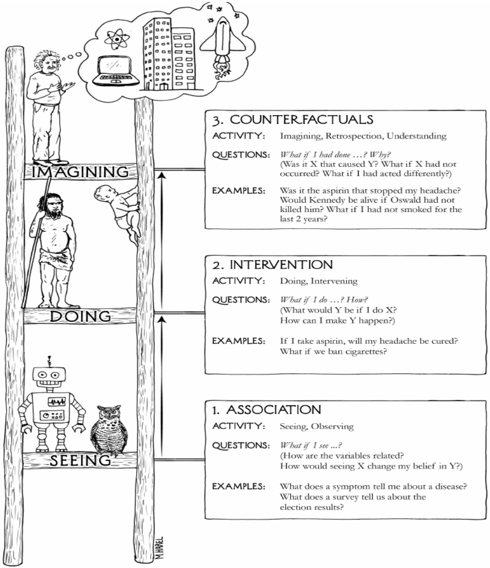
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Intervention

- Moving beyond observation by introducing actions to change the system and observing the results: "What if I do this?"
- Represent causal relationship by showing how changes in one variable affect another
- Question: "What will happen to Y if I change X?"
- Methods: randomized controlled trials, where the experimenter manipulates the independent variable $P(Y|do(X))$ (e.g. giving patients a treatment placebo) to observe the effect on a dependent variable (like recovery rates)



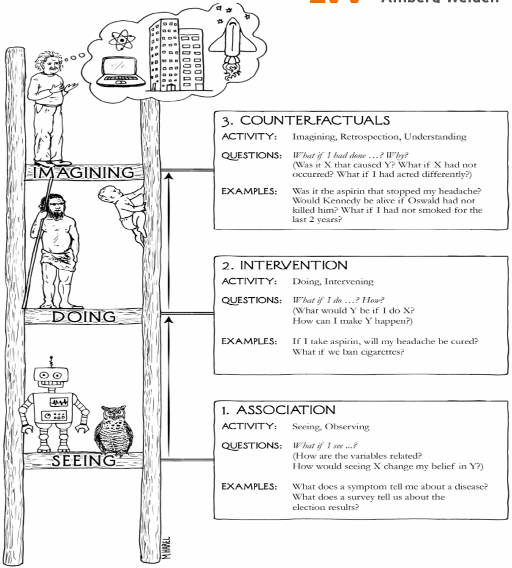
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Counterfactuals

- Asking about hypothetical alternatives to what actually happened, what would have happened "if"



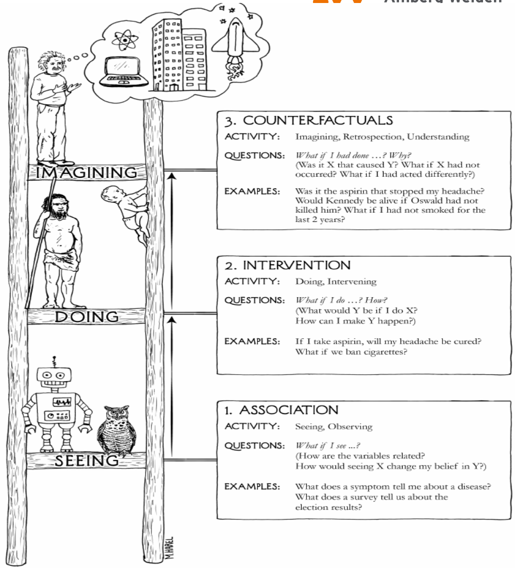
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Counterfactuals

- Asking about hypothetical alternatives to what actually happened, what would have happened "if"
- Alternative realities by imagining what the outcome would have been if a different action or condition had been in place



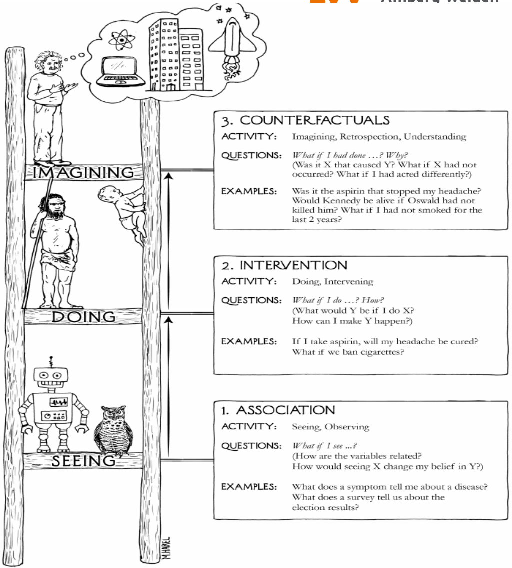
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

Counterfactuals

- Asking about hypothetical alternatives to what actually happened, what would have happened "if"
- Alternative realities by imagining what the outcome would have been if a different action or condition had been in place
- Question: "What would have happened to Y if X had been different?"



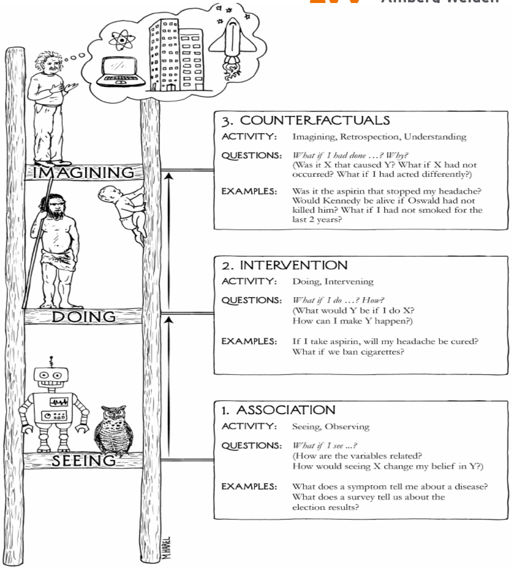
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

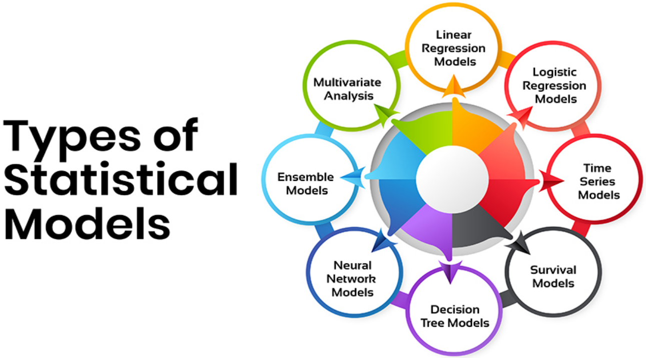
Causal Learning

Counterfactuals

- Asking about hypothetical alternatives to what actually happened, what would have happened "if"
- Alternative realities by imagining what the outcome would have been if a different action or condition had been in place
- Question: "What would have happened to Y if X had been different?"
- Methods: structural causal models (SCMs), provide a framework for understanding causal relationships at a deeper level by specifying equations reflecting mechanisms of causation



Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

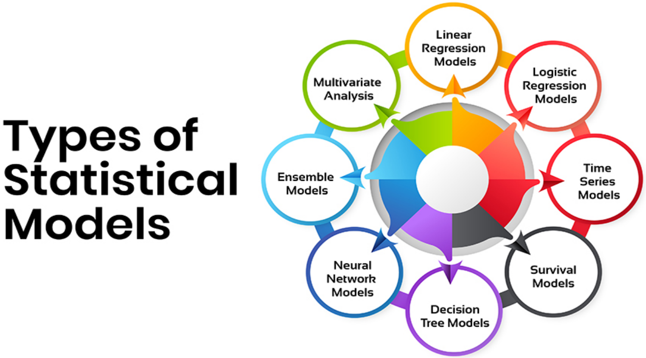


- Current ML-Models or statistical models can only do "Association"

Source: https://www.dasca.org/world-of-data-science/article/what-is-statistical-modeling-in-data-science
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

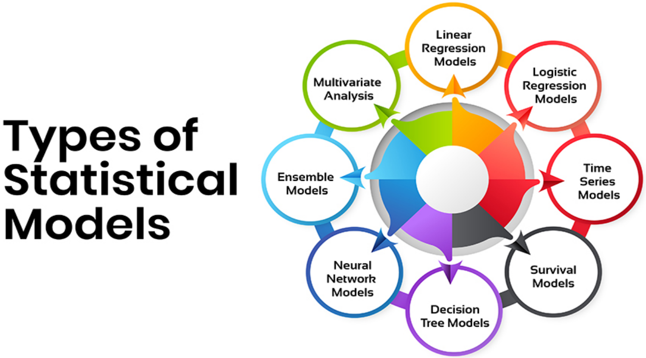


- Current ML-Models or statistical models can only do “Association”
- Statistical inference $\rightarrow P(Fire|Smoke), P(Smoke|Fire)$

Source: https://www.dasca.org/world-of-data-science/article/what-is-statistical-modeling-in-data-science
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

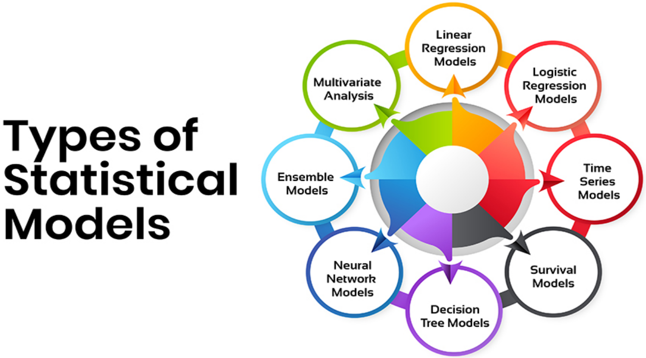
Deep Learning Strategies – Part III

Causal Learning



- Current ML-Models or statistical models can only do "Association"
- Statistical inference $\rightarrow$ $P(Fire|Smoke)$, $P(Smoke|Fire)$
- No Causal Inference $\rightarrow$ Yes: Fire causes Smoke, No: Smoke causes Fire

Source: https://www.dasca.org/world-of-data-science/article/what-is-statistical-modeling-in-data-science
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning



- Current ML-Models or statistical models can only do "Association"
- Statistical inference → $P(Fire|Smoke)$, $P(Smoke|Fire)$
- No Causal Inference → Yes: Fire causes Smoke, No: Smoke causes Fire
- Successful because: massive amount of data, high number of parameters, strong computational power, independent and identically distributed (i.i.d.) assumption

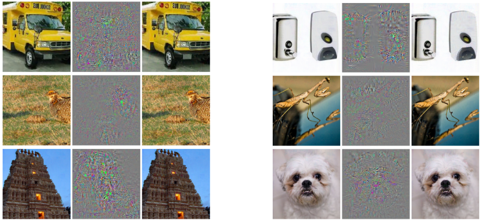
Source: https://www.dasca.org/world-of-data-science/article/what-is-statistical-modeling-in-data-science

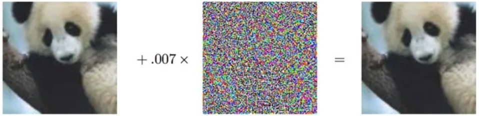
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

- Independent: each data sample is collected independently of others, one sample does not influence/depend on another





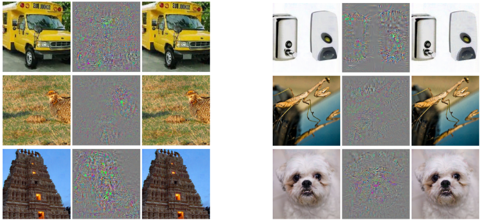
Source: Szegedy et al., Intriguing properties of neural networks, Figure 5

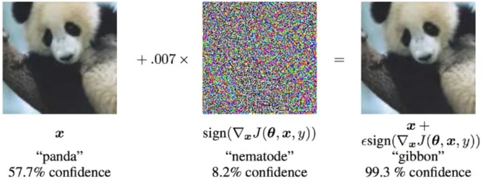
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

- Independent: each data sample is collected independently of others, one sample does not influence/depend on another
- Identically Distributed: all samples from the same probability distribution → Same data distribution during training as during testing or deployment





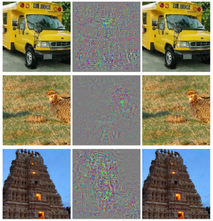
Source: Szegedy et al., Intriguing properties of neural networks, Figure 5

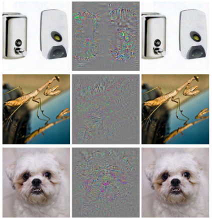
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

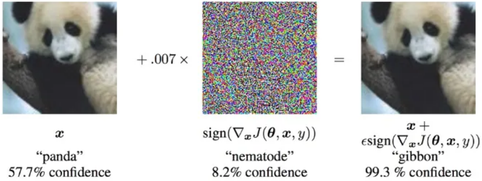
Deep Learning Strategies – Part III

Causal Learning

- Independent: each data sample is collected independently of others, one sample does not influence/depend on another
- Identically Distributed: all samples from the same probability distribution → Same data distribution during training as during testing or deployment
- Real-world Distribution Shifts: changes in the underlying data distribution between training, testing or deployment:
  - Temporal: properties evolve over time
  - Contextual: data in one environment/setup not representative in another, different devices, methods, sources







Source: Szegedy et al., Intriguing properties of neural networks, Figure 5

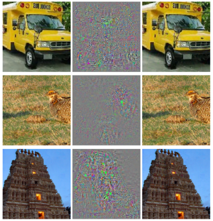
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

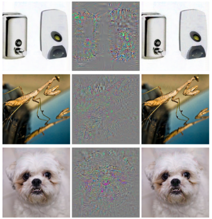
Deep Learning Strategies – Part III

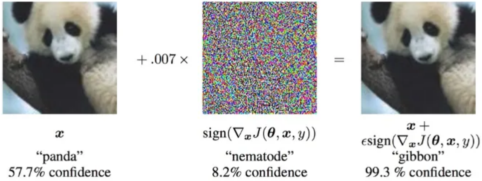
OTH Amberg-Weiden

Causal Learning

- Independent: each data sample is collected independently of others, one sample does not influence/depend on another
- Identically Distributed: all samples from the same probability distribution $\rightarrow$ Same data distribution during training as during testing or deployment
- Real-world Distribution Shifts: changes in the underlying data distribution between training, testing or deployment:
  - Temporal: properties evolve over time
  - Contextual: data in one environment/setup not representative in another, different devices, methods, sources
- Introducing an unexpected distribution change?







Source: Szegedy et al., Intriguing properties of neural networks, Figure 5

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

OTH Amberg-Weiden

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data
- Help towards important goals of ML: Generalization, Transfer Learning, Explainability

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

OTH Amberg-Weiden

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data
- Help towards important goals of ML: Generalization, Transfer Learning, Explainability
- Explains the connection between causality and statistical dependence

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data
- Help towards important goals of ML: Generalization, Transfer Learning, Explainability
- Explains the connection between causality and statistical dependence
- Postulated by Hans Reichenbach in 1956: given two observations X and Y are statistically dependent $\rightarrow P(X, Y) \neq P(X)P(Y)$, then an event Z which causally influences both results in $P(X, Y|Z) = P(X|Z)P(Y|Z)$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data
- Help towards important goals of ML: Generalization, Transfer Learning, Explainability
- Explains the connection between causality and statistical dependence
- Postulated by Hans Reichenbach in 1956: given two observations X and Y are statistically dependent $\rightarrow P(X, Y) \neq P(X)P(Y)$, then an event Z which causally influences both results in $P(X, Y|Z) = P(X|Z)P(Y|Z)$
- If Z is not observed (hidden), it is a confounding variable, affecting both the cause and the effect, making it difficult to establish a true causal relationship

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

OTH Amberg-Weiden

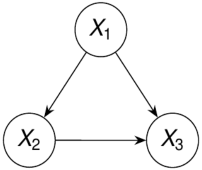
Common Causal Principles

- Causal Learning will help ML-models to plan and act according to interventions
- Robust to distribution shifts, using counterfactual questions to learn unseen data
- Help towards important goals of ML: Generalization, Transfer Learning, Explainability
- Explains the connection between causality and statistical dependence
- Postulated by Hans Reichenbach in 1956: given two observations X and Y are statistically dependent $\rightarrow P(X, Y) \neq P(X)P(Y)$, then an event Z which causally influences both results in $P(X, Y|Z) = P(X|Z)P(Y|Z)$
- If Z is not observed (hidden), it is a confounding variable, affecting both the cause and the effect, making it difficult to establish a true causal relationship
- Example: exercise frequency (Z), body weight (X), cholesterol levels (Y) $\rightarrow$ without considering Z, it might appear that body weight directly affects cholesterol, however, exercise impacts both independently

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Causal Graphical Models (CGMs)

- Directed Acyclic Graphs (DAGs) or Bayesian Networks ("acyclic" part $\rightarrow$ no loops, flow in one direction without cycling back)

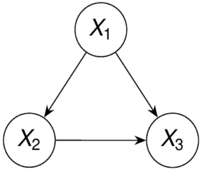


$$p(X_1, X_2, X_3) = p(X_1)\, p(X_2 \mid X_1)\, p(X_3 \mid X_1, X_2)$$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Causal Graphical Models (CGMs)

- Directed Acyclic Graphs (DAGs) or Bayesian Networks ("acyclic" part $\rightarrow$ no loops, flow in one direction without cycling back)
- Causal relationships of variables using graph structures

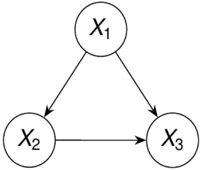


$$p(X_1, X_2, X_3) = p(X_1)\, p(X_2 \mid X_1)\, p(X_3 \mid X_1, X_2)$$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Causal Graphical Models (CGMs)

- Directed Acyclic Graphs (DAGs) or Bayesian Networks (“acyclic” part $\rightarrow$ no loops, flow in one direction without cycling back)
- Causal relationships of variables using graph structures
- How different variables in a system influence each other, clarifying potential cause-and-effect paths



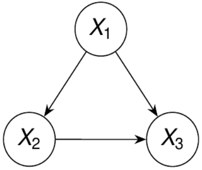
$$p(X_1, X_2, X_3) = p(X_1)\, p(X_2 \mid X_1)\, p(X_3 \mid X_1, X_2)$$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

OTH Amberg-Weiden

Causal Graphical Models (CGMs)

- Directed Acyclic Graphs (DAGs) or Bayesian Networks ("acyclic" part $\rightarrow$ no loops, flow in one direction without cycling back)
- Causal relationships of variables using graph structure:
- How different variables in a system influence each other, clarifying potential cause-and-effect paths
- Variables X (observations) as nodes and causal relationships between them as directed edges



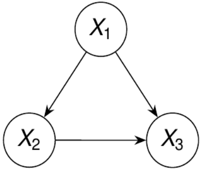
$$p(X_1, X_2, X_3) = p(X_1)\, p(X_2 \mid X_1)\, p(X_3 \mid X_1, X_2)$$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

OTH Amberg-Weiden

Causal Graphical Models (CGMs)

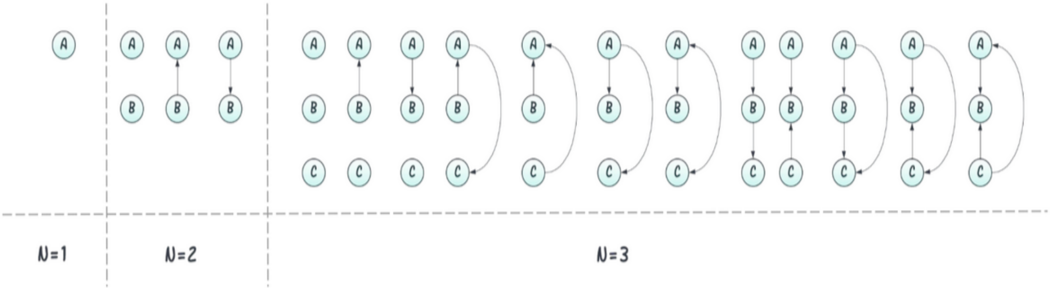
- Directed Acyclic Graphs (DAGs) or Bayesian Networks ("acyclic" part $\rightarrow$ no loops, flow in one direction without cycling back)
- Causal relationships of variables using graph structures
- How different variables in a system influence each other, clarifying potential cause-and-effect paths
- Variables X (observations) as nodes and causal relationships between them as directed edges
- Causal Markov condition (Causal Independence!): $P(X_1, ..., X_n) = \prod_{i=1}^{n} P(X_i | PA_i)$, with PA_i as the parents of X_i, having a direct causal link (e.g. $X_1 \rightarrow X_2$)



$$p(X_1, X_2, X_3) = p(X_1)\, p(X_2 \mid X_1)\, p(X_3 \mid X_1, X_2)$$

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

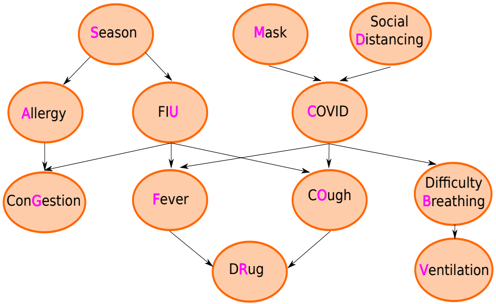
Space of causal graphs (i.e. DAGs) grows **super-exponentially**.



Source: Image from Nan Rosemary Ke, Stefan Bauer, "Causality and Deep Learning: Synergies, Challenges and the Future", Slide 26

Bayesian Network

#	COVID	Mask	Soc. Distance	FIU	Cough	Fever	Ventilate	Season	ConGestion	Difficulty Breath	Drug	Allergy
1	1	0	1	0	1	0	1	Spring	0	0	1	0
2	0	1	0	1	0	1	0	Summer	1	1	0	0
3	1	1	1	0	0	0	1	Fall	1	0	1	1
4	0	0	1	1	1	0	0	Winter	0	1	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...
500	0	1	1	1	0	0	0	Summer	0	0	1	1

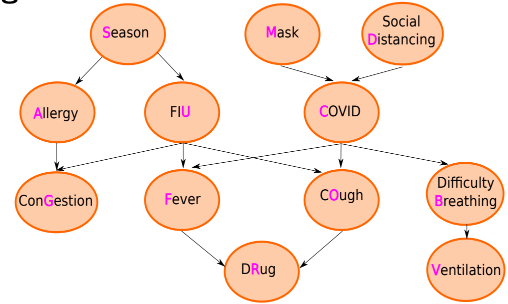


Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Deep Learning Strategies – Part III

Causal Learning

Bayesian Network



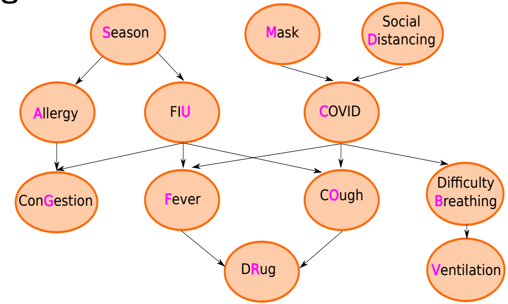
- Random (dependent/independent) variables, e.g. Fever/COVID and Allergy/Cough

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Deep Learning Strategies – Part III

Causal Learning

Bayesian Network

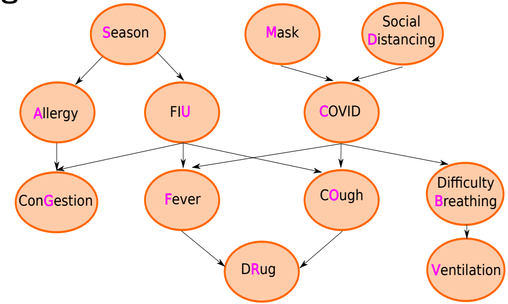


- Random (dependent/independent) variables, e.g. Fever/COVID and Allergy/Cough
- Joint (dependent) probability: from COVID $\rightarrow$ Fever – $P(F, C) = P(F|C)P(C)$ – or from Fever $\rightarrow$ COVID – $P(F, C) = P(C|F)P(F)$

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

OTH
Amberg-Weiden

Bayesian Network



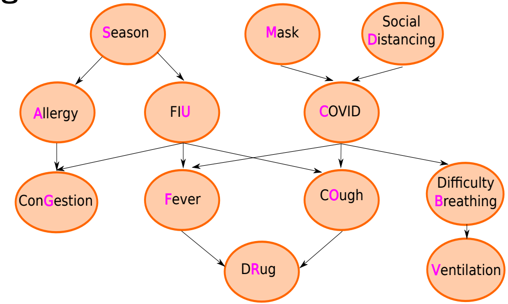
- Random (dependent/independent) variables, e.g. Fever/COVID and Allergy/Cough
- Joint (dependent) probability: from COVID $\rightarrow$ Fever – $P(F, C) = P(F|C)P(C)$ – or from Fever $\rightarrow$ COVID – $P(F, C) = P(C|F)P(F)$
- Valid for describing Bayesian Network, however: Links do not describe Causality!

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Deep Learning Strategies – Part III

Causal Learning

Bayesian Network

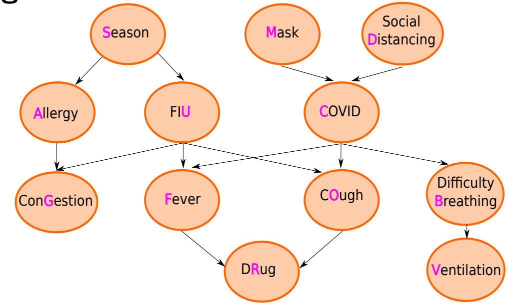


- Random (dependent/independent) variables, e.g. Fever/COVID and Allergy/Cough
- Joint (dependent) probability: from COVID $\rightarrow$ Fever – $P(F, C) = P(F|C)P(C)$ – or from Fever $\rightarrow$ COVID – $P(F, C) = P(C|F)P(F)$
- Valid for describing Bayesian Network, however: Links do not describe Causality!
- Causality (single direction!) would mean either F is causing C or C is causing F

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

OTH Amberg-Weiden

Bayesian Network



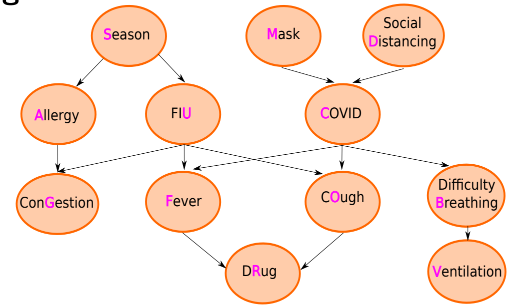
- Random (dependent/independent) variables, e.g. Fever/COVID and Allergy/Cough
- Joint (dependent) probability: from COVID → Fever – $P(F, C) = P(F|C)P(C)$ – or from Fever → COVID – $P(F, C) = P(C|F)P(F)$
- Valid for describing Bayesian Network, however: Links do not describe Causality!
- Causality (single direction!) would mean either F is causing C or C is causing F
- How to read statistical dependence as causal dependence?

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Deep Learning Strategies – Part III

Causal Learning

Bayesian Network



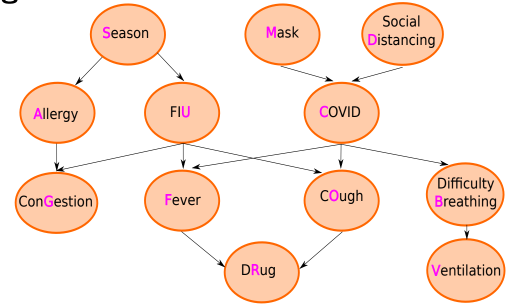
- In general: read a conditional probability $P(X|Y_1, ..., Y_n) \neq P(X)$, as X is "caused" by $Y_1, ..., Y_n$

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Deep Learning Strategies – Part III

Causal Learning

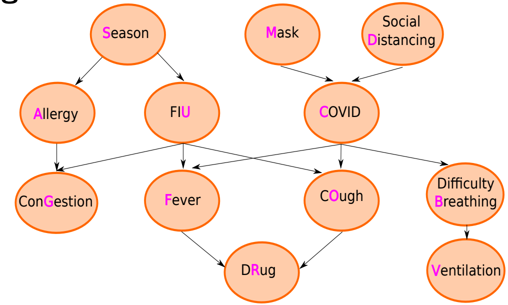
Bayesian Network



- In general: read a conditional probability $P(X|Y_1, \ldots, Y_n) \neq P(X)$, as X is "caused" by $Y_1, \ldots, Y_n$
- However: above relation just indicates association, meaning how X is associated with $Y_1, \ldots, Y_n$, also possible: $P(Y_1, \ldots, Y_n|X) \neq P(Y_1, \ldots, Y_n)$

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

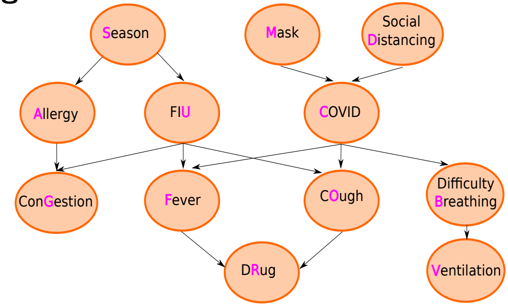
Bayesian Network



- In general: read a conditional probability $P(X|Y_1, ..., Y_n) \neq P(X)$, as X is "caused" by $Y_1, ..., Y_n$
- However: above relation just indicates association, meaning how X is associated with $Y_1, ..., Y_n$, also possible: $P(Y_1, ..., Y_n|X) \neq P(Y_1, ..., Y_n)$
- Still: No Causality!

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

Bayesian Network

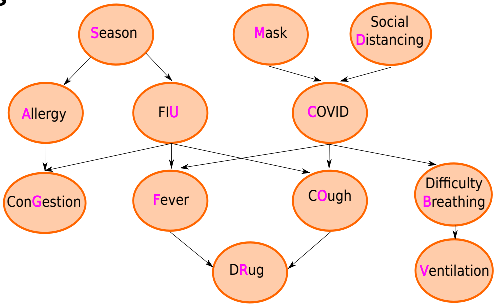


- In general: read a conditional probability $P(X|Y_1, ..., Y_n) \neq P(X)$, as X is "caused" by $Y_1, ..., Y_n$
- However: above relation just indicates association, meaning how X is associated with $Y_1, ..., Y_n$, also possible: $P(Y_1, ..., Y_n|X) \neq P(Y_1, ..., Y_n)$
- Still: No Causality!
- Goal: Given data D, what is causing what (the other)?

Source: YouTube – easy learning – Causality 1: Bayesian networks are not causal, Link

OTH Amberg-Weiden

Bayesian Network



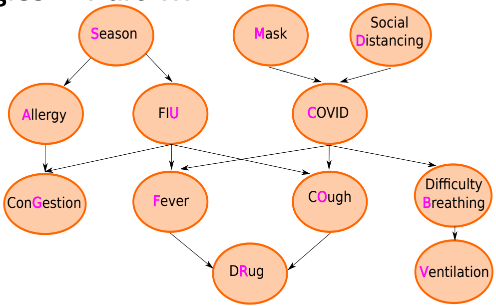
- Consider binary variables COVID $\in$ {COVID, –COVID} and Fever $\in$ {Fever, –Fever}

Source: YouTube – easy learning – Causality 2: Intervention, Link

Deep Learning Strategies – Part III

Causal Learning

Bayesian Network



- Consider binary variables COVID $\in$ {COVID, –COVID} and Fever $\in$ {Fever, –Fever}
- COVID is detected by a test, Fever by temperature, while observations result in:

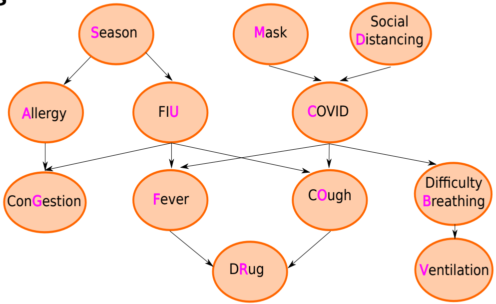
$$P(\text{Fever}|\text{COVID}) > P(\text{Fever}|\text{–COVID})$$

$$P(\text{COVID}|\text{Fever}) > P(\text{COVID}|\text{–Fever})$$

Source: YouTube – easy learning – Causality 2: Intervention, Link

Causal Learning

Bayesian Network



- Consider binary variables COVID $\in$ {COVID, –COVID} and Fever $\in$ {Fever, –Fever}
- COVID is detected by a test, Fever by temperature, while observations result in:

$$P(\text{Fever}|\text{COVID}) > P(\text{Fever}|\text{–COVID})$$

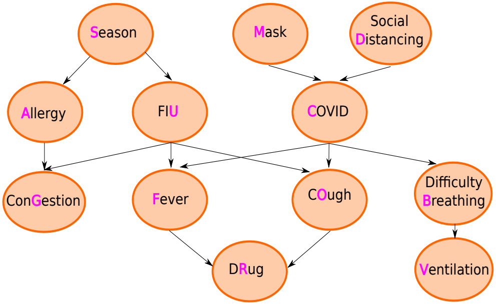
$$P(\text{COVID}|\text{Fever}) > P(\text{COVID}|\text{–Fever})$$

- Observing one of them, increases probability of the other! $\rightarrow$ Statistical dependence (correlation!) – Which one causes the other, how to figure it out – more data?

Source: YouTube – easy learning – Causality 2: Intervention, Link

OTH Amberg-Weiden

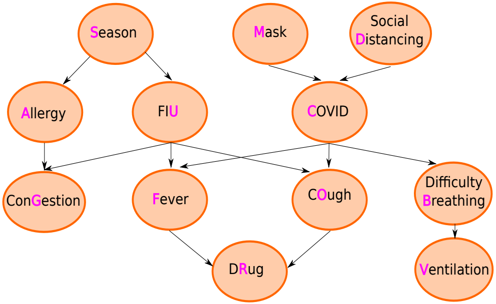
Observation vs. Intervention



- Just by observing, it is not possible to determine what is causality!

Source: YouTube – easy learning – Causality 2: Intervention, Link

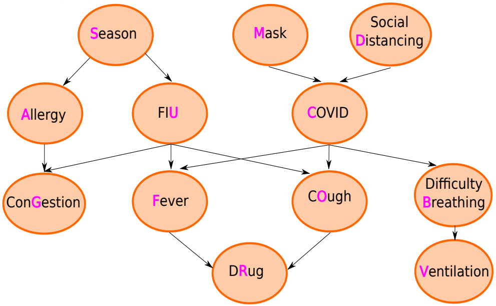
Observation vs. Intervention



- Just by observing, it is not possible to determine what is causality!
- Why??? – How do we know that COVID is causing ($\rightarrow$) Fever and not vice versa?

Source: YouTube – easy learning – Causality 2: Intervention, Link

Observation vs. Intervention

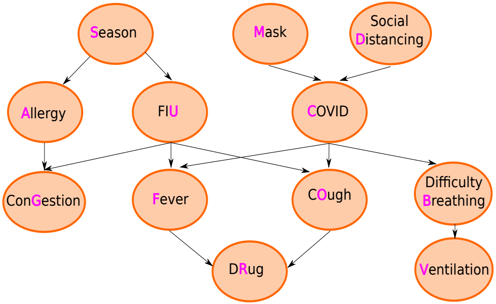


- Just by observing, it is not possible to determine what is causality!
- Why??? – How do we know that COVID is causing ($\rightarrow$) Fever and not vice versa?
- Intervene via experimental verification!!! (by observing people with COVID, it is likely to observe Fever; observing Fever, is very often not COVID-related, e.g. fire/smoke)

Source: YouTube – easy learning – Causality 2: Intervention, Link

OTH Amberg-Weiden

Observation vs. Intervention



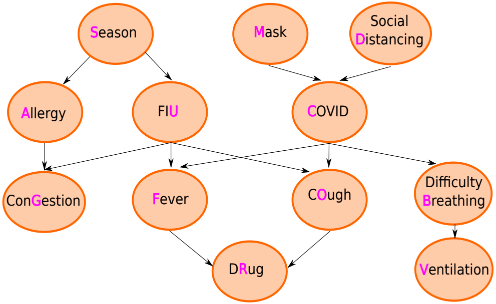
- Just by observing, it is not possible to determine what is causality!
- Why??? – How do we know that COVID is causing ($\rightarrow$) Fever and not vice versa?
- Intervene via experimental verification!!! (by observing people with COVID, it is likely to observe Fever; observing Fever, is very often not COVID-related, e.g. fire/smoke)
- Force variables to certain values and check changes in the other variable

Source: YouTube – easy learning – Causality 2: Intervention, Link

Deep Learning Strategies – Part III

Causal Learning

Observation vs. Intervention



- Just by observing, it is not possible to determine what is causality!
- Why??? – How do we know that COVID is causing ($\rightarrow$) Fever and not vice versa?
- Intervene via experimental verification!!! (by observing people with COVID, it is likely to observe Fever; observing Fever, is very often not COVID-related, e.g. fire/smoke)
- Force variables to certain values and check changes in the other variable
- Intervention $\rightarrow$ Causality!!! (...Find causal relationships)

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\mathrm{do}(Z := z)$ or $\mathrm{do}(z)$

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\mathrm{do}(Z := z)$ or $\mathrm{do}(z)$
- $P(Y|\mathrm{do}(Z := z)) \rightarrow$ Probability of Y, given that Z is set/forced to z

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\mathrm{do}(Z := z)$ or $\mathrm{do}(z)$
- $P(Y|\mathrm{do}(Z := z)) \rightarrow$ Probability of Y, given that Z is set/forced to z
- $P(Y|Z = z) \rightarrow$ Probability of Y, given that Z is seen/observed equals to z

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\mathrm{do}(Z := z)$ or $\mathrm{do}(z)$
- $P(Y|\mathrm{do}(Z := z)) \rightarrow$ Probability of Y, given that Z is set/forced to z
- $P(Y|Z = z) \rightarrow$ Probability of Y, given that Z is seen/observed equals to z
- Constraint: when intervening on one variable, no other variable should be intervened

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\text{do}(Z := z)$ or $\text{do}(z)$
- $P(Y|\text{do}(Z := z)) \rightarrow$ Probability of Y, given that Z is set/forced to z
- $P(Y|Z = z) \rightarrow$ Probability of Y, given that Z is seen/observed equals to z
- Constraint: when intervening on one variable, no other variable should be intervened
- $\text{do}(\text{COVID} := 1)$ or $\text{do}(\text{COVID} := 0) \rightarrow$ Expectation: affect the likelihood of Fever, because COVID is a cause of fever (being the effect/sympotme of COVID)

Source: YouTube – easy learning – Causality 2: Intervention, Link

Intervention – Do-Operator

- Definition "Intervention": an ideal intervention on a given random variable Z is described by forcing its value to z, denoted $\mathrm{do}(Z := z)$ or $\mathrm{do}(z)$
- $P(Y|\mathrm{do}(Z := z)) \rightarrow$ Probability of Y, given that Z is set/forced to z
- $P(Y|Z = z) \rightarrow$ Probability of Y, given that Z is seen/observed equals to z
- Constraint: when intervening on one variable, no other variable should be intervened
- $\mathrm{do}(\mathrm{COVID} := 1)$ or $\mathrm{do}(\mathrm{COVID} := 0) \rightarrow$ Expectation: affect the likelihood of Fever, because COVID is a cause of fever (being the effect/sympotme of COVID)
- $\mathrm{do}(\mathrm{Fever} := 1)$ or $\mathrm{do}(\mathrm{Fever} := 0) \rightarrow$ Expectation: does not affect the likelihood of COVID, because Fever is NOT a cause of COVID (not a sympotme of Fever)

Source: YouTube – easy learning – Causality 2: Intervention, Link

- Intervene on Fever (induce/remove fever) and observe COVID (person has/has not COVID), should not result in any probability difference of seeing/causing COVID:

$$P(\text{COVID}|\text{do}(\text{Fever})) = P(\text{COVID}|\text{do}(\text{–Fever}))$$

indicating that Fever does not cause COVID!

$$P(\text{COVID}|\text{see}(\text{Fever})) > P(\text{COVID}|\text{see}(\text{–Fever}))$$

Different from the observational relation!

Source: YouTube – easy learning – Causality 2: Intervention, Link

- Intervene on Fever (induce/remove fever) and observe COVID (person has/has not COVID), should not result in any probability difference of seeing/causing COVID:

$$P(\text{COVID}|\text{do}(\text{Fever})) = P(\text{COVID}|\text{do}(-\text{Fever}))$$

  indicating that Fever does not cause COVID!

$$P(\text{COVID}|\text{see}(\text{Fever})) > P(\text{COVID}|\text{see}(-\text{Fever}))$$

  Different from the observational relation!

- However: intervene on COVID and observe Fever, lead to:

$$P(\text{Fever}|\text{do}(\text{COVID})) > P(\text{Fever}|\text{do}(-\text{COVID}))$$

  COVID causes Fever!

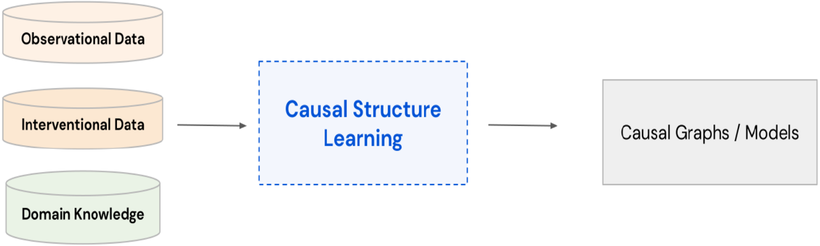
$$P(\text{Fever}|\text{COVID}) > P(\text{Fever}|-\text{COVID})$$

  Interventional and observational relation is in line!

Source: YouTube – easy learning – Causality 2: Intervention, Link

Deep Learning Strategies – Part III

Causal Learning

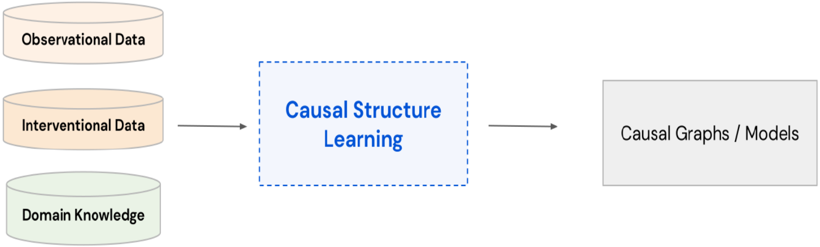


- Observational Data: collected without actively manipulating variables, just observing the natural state of the system (correlations/associations)

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning

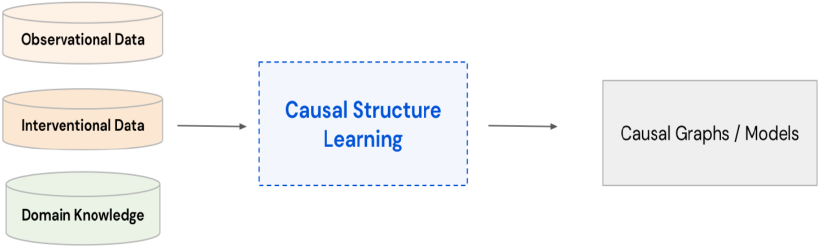


- Observational Data: collected without actively manipulating variables, just observing the natural state of the system (correlations/associations)
- Interventional Data: collected by actively changing or intervening on specific variables, to reveal causal relationships

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Deep Learning Strategies – Part III

Causal Learning



- Observational Data: collected without actively manipulating variables, just observing the natural state of the system (correlations/associations)
- Interventional Data: collected by actively changing or intervening on specific variables, to reveal causal relationships
- Domain Knowledge: includes expert understanding, assumptions, or prior information about the system being studied, to help guiding the interpretation of observational data and the design of meaningful interventions

Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

- Average Causal Effect (ACE): Binary variable $Z \in \{0, 1\}$, causes variable Y, if the average causal effect (ACE) of Z on $Y = y$ is defined as:

$$ACE(Z \to Y_y) = P(Y = y|\text{do}(Z := 1)) - P(Y = y|\text{do}(Z := 0))$$

is non-zero!

Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

- Average Causal Effect (ACE): Binary variable $Z \in \{0, 1\}$, causes variable Y, if the average causal effect (ACE) of Z on $Y = y$ is defined as:

$$ACE(Z \rightarrow Y_y) = P(Y = y|\text{do}(Z := 1)) - P(Y = y|\text{do}(Z := 0))$$

  is non-zero!

- Causality: A variable Z causes variable Y if there exist two interventional events z_0 and z_1 changing the distribution Y:

$$P(Y|\text{do}(z_0)) \neq P(Y|\text{do}(z_1))$$

Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

- Average Causal Effect (ACE): Binary variable $Z \in \{0, 1\}$, causes variable Y, if the average causal effect (ACE) of Z on $Y = y$ is defined as:

$$ACE(Z \to Y_y) = P(Y = y|\text{do}(Z := 1)) - P(Y = y|\text{do}(Z := 0))$$

  is non-zero!

- Causality: A variable Z causes variable Y if there exist two interventional events z_0 and z_1 changing the distribution Y:

$$P(Y|\text{do}(z_0)) \neq P(Y|\text{do}(z_1))$$

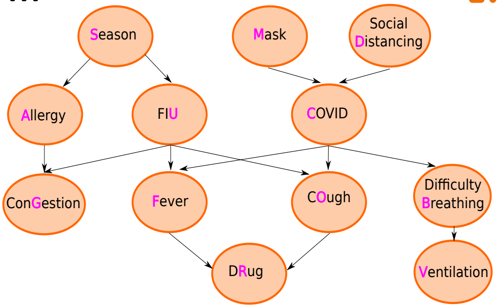
- However: Let $Y = X \oplus Z$ be the XOR-Function of X and Z, then:

$$P(Y|\text{do}(X = 0)) = P(Y|\text{do}(X = 1)) = P(Y)$$

  $\rightarrow$ Better Definition as Deterministic Function!

Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Causality as Deterministic Function!

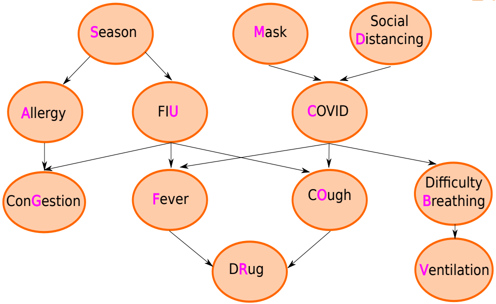


- COVID is the cause of Cough and Cough is the cause of Drug → More factors, however not being explicitly modeled and represented as single exogenous (latent variable) U_X

Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

OTH Amberg-Weiden

Causality as Deterministic Function!

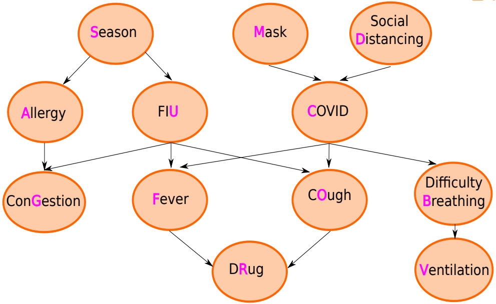


- COVID is the cause of Cough and Cough is the cause of Drug → More factors, however not being explicitly modeled and represented as single exogenous (latent variable) U_X
- Each Variable X (here: C, O, R – referred to as endogenous) have such exogenous (here: U_C, U_O, U_R) variables, used together with a deterministic function of is endogenous and exogenous parents

Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

OTH Amberg-Weiden

Causality as Deterministic Function!



- COVID is the cause of Cough and Cough is the cause of Drug $\rightarrow$ More factors, however not being explicitly modeled and represented as single exogenous (latent variable) U_X
- Each Variable X (here: C, O, R – referred to as endogenous) have such exogenous (here: U_C, U_O, U_R) variables, used together with a deterministic function of is endogenous and exogenous parents
- $C := f_C(U_C) = U_C$, $O := f_O(C, U_O)$, $R := f_R(O, U_R)$
  $\rightarrow$ Structural Causal Model (SCM)

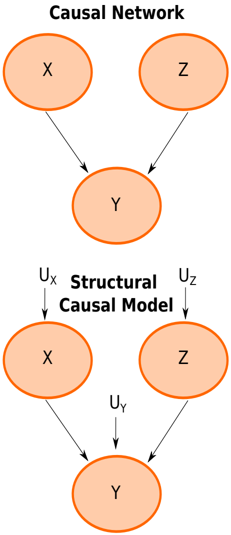
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs)

- SCM defined as a touple $< V, U, P_U, F >$ with:

1. V as the set of endogenous random variables
2. $U = \{U_X | X \in V\}$ as set of exogenous/latent/noise random variables
3. P_U is the joint distribution over U satisfying $P_U(U) = \prod_X P_U(U_X)$
4. F set of functions f_X, how values are assigned to each variable $X \in V$, based on a subset of endogenous variables $Pa_X \subset V \setminus \{X\}$ (parents), with a parent exogenous variable $U_X \subset U$, leading to:

$$X := f_X(Pa_X, U_X)$$



Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

OTH Amberg-Weiden

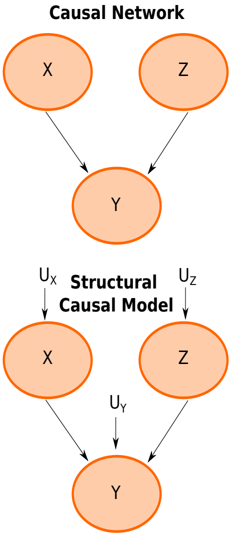
Structural Causal Models (SCMs)

- SCM defined as a touple $< V, U, P_U, F >$ with:

1. V as the set of endogenous random variables
2. $U = \{U_X | X \in V\}$ as set of exogenous/latent/noise random variables
3. P_U is the joint distribution over U satisfying $P_U(U) = \prod_X P_U(U_X)$
4. F set of functions f_X, how values are assigned to each variable $X \in V$, based on a subset of endogenous variables $Pa_X \subset V \setminus \{X\}$ (parents), with a parent exogenous variable $U_X \subset U$, leading to:

$$X := f_X(Pa_X, U_X)$$

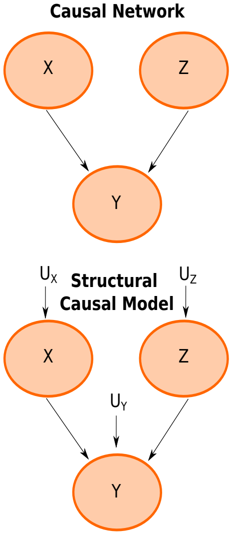
- Causal graph (network) over nodes V, where the parents of each node are linked to it (assume the graph as DAG)



Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs)

- Parents of a variable V are also called direct causes (one hierarchy level above), while the ancestors are known as indirect causes (multiple previous hierarchy levels)

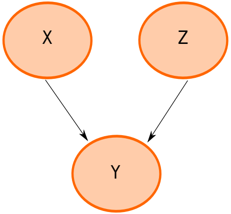


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

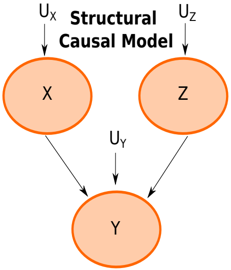
Structural Causal Models (SCMs)

- Parents of a variable V are also called direct causes (one hierarchy level above), while the ancestors are known as indirect causes (multiple previous hierarchy levels)
- In that specific case both variables X and Z do not possess endogenous parents, however, only exogenous parents U_X, U_Z

Causal Network

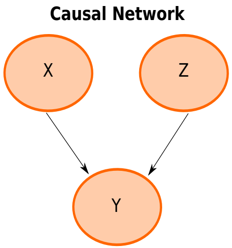


Structural Causal Model

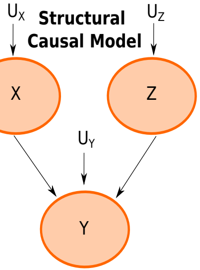


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs)



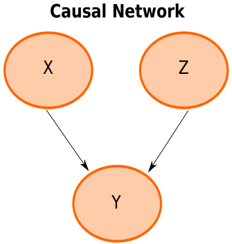
- Parents of a variable V are also called direct causes (one hierarchy level above), while the ancestors are known as indirect causes (multiple previous hierarchy levels)
- In that specific case both variables X and Z do not possess endogenous parents, however, only exogenous parents U_X, U_Z
- The variable $X = f_X(Pa_X, U_X)$ (e.g. Y) depends only on endogenous parents (e.g X, Z), together with the exogenous variable (e.g U_Y)

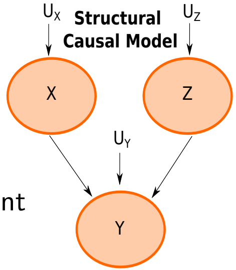


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs)

- Parents of a variable V are also called direct causes (one hierarchy level above), while the ancestors are known as indirect causes (multiple previous hierarchy levels)
- In that specific case both variables X and Z do not possess endogenous parents, however, only exogenous parents U_X, U_Z
- The variable $X = f_X(Pa_X, U_X)$ (e.g. Y) depends only on endogenous parents (e.g X, Z), together with the exogenous variable (e.g U_Y)
- Observational distribution of all endogenous variables V refers to the joint probability distribution $P(V) = \prod_X P(X|Pa_X)$

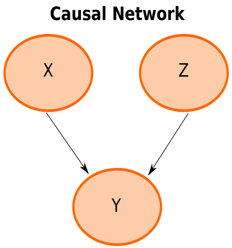


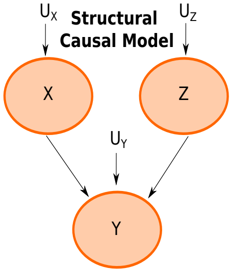


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs)

- The causal network $\mathcal{G}$ is an I-Map for $P(V) = \prod_X P(X|Pa_X)$ (Bayesian Network Factorization) with $P(X|Pa_X)$ obtained from $X = f_X(Pa_X, U_X)$

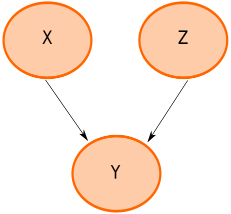




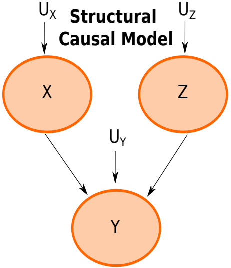
Structural Causal Models (SCMs)

- The causal network $\mathcal{G}$ is an I-Map for $P(V) = \prod_X P(X|Pa_X)$ (Bayesian Network Factorization) with $P(X|Pa_X)$ obtained from $X = f_X(Pa_X, U_X)$
- I-MAP (Independence Map) for the probability distribution P: graph (usually DAG) of a set of variables, capturing conditional independence

Causal Network

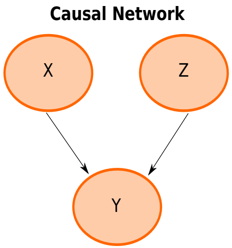


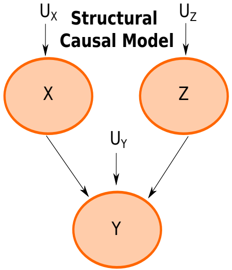
Structural Causal Model



Structural Causal Models (SCMs)

- The causal network $\mathcal{G}$ is an I-Map for $P(V) = \prod_X P(X|Pa_X)$ (Bayesian Network Factorization) with $P(X|Pa_X)$ obtained from $X = f_X(Pa_X, U_X)$
- I-MAP (Independence Map) for the probability distribution P: graph (usually DAG) of a set of variables, capturing conditional independence
- Conditional independence: situation where two variables X & Z, are independent given a third variable Y, leading to $P(X, Z|Y) = P(X|Y)P(Z|Y)$

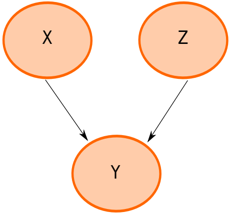




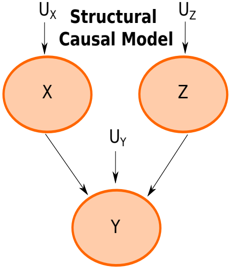
Structural Causal Models (SCMs)

- The causal network $\mathcal{G}$ is an I-Map for $P(V) = \prod_X P(X|Pa_X)$ (Bayesian Network Factorization) with $P(X|Pa_X)$ obtained from $X = f_X(Pa_X, U_X)$
- I-MAP (Independence Map) for the probability distribution P: graph (usually DAG) of a set of variables, capturing conditional independence
- Conditional independence: situation where two variables X & Z, are independent given a third variable Y, leading to $P(X, Z|Y) = P(X|Y)P(Z|Y)$
- Example I-MAP Graph: $X \rightarrow Y \rightarrow Z$ – with $X =$ Having Disease, $Y =$ Taking Medication, $Z =$ Observing Side Effects

Causal Network



Structural Causal Model

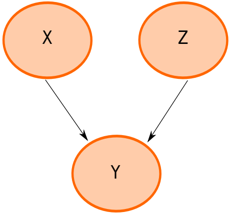


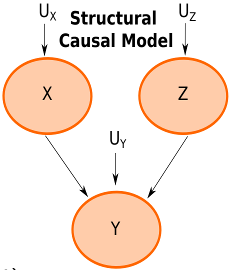
Deep Learning Strategies – Part III
Causal Learning

Structural Causal Models (SCMs)

- The causal network $\mathcal{G}$ is an I-Map for $P(V) = \prod_X P(X|Pa_X)$ (Bayesian Network Factorization) with $P(X|Pa_X)$ obtained from $X = f_X(Pa_X, U_X)$
- I-MAP (Independence Map) for the probability distribution P: graph (usually DAG) of a set of variables, capturing conditional independence
- Conditional independence: situation where two variables X & Z, are independent given a third variable Y, leading to $P(X, Z|Y) = P(X|Y)P(Z|Y)$
- Example I-MAP Graph: $X \rightarrow Y \rightarrow Z$ – with $X =$ Having Disease, $Y =$ Taking Medication, $Z =$ Observing Side Effects
- Captures direct dependencies ($X \rightarrow Y$, $Y \rightarrow Z$) and conditional independencies (X and Z are independent given Y, disease X alone not directly lead to side effects Z, only through the medication Y – parents!!!)

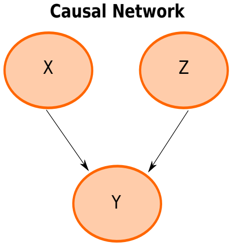
Causal Network

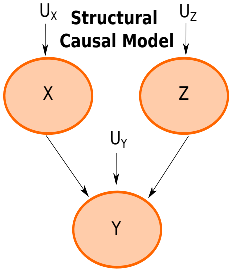




Structural Causal Models (SCMs) – Example XOR

- $P(V) = P(X = x, Z = z, Y = y)$

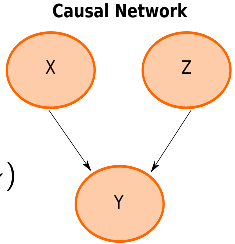


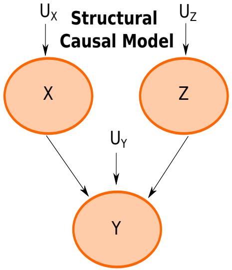


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR

- $P(V) = P(X = x, Z = z, Y = y)$
- $P(V) = \prod_X P(X|Pa_X) = P(X = x|Pa_X) \cdot P(Z = z|Pa_Z) \cdot P(Y = y|Pa_Y)$



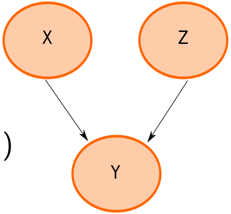


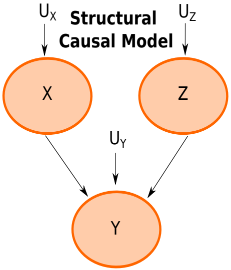
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR

- $P(V) = P(X = x, Z = z, Y = y)$
- $P(V) = \prod_X P(X|Pa_X) = P(X = x|Pa_X) \cdot P(Z = z|Pa_Z) \cdot P(Y = y|Pa_Y)$
- $X = f_X(Pa_X, U_X) = U_x$
- $Z = f_Z(Pa_Z, U_z) = U_z$
- $Y = f_Y(Pa_Y, U_Y)$, with $U_Y = 0 \rightarrow Y = (X, Z)$

Causal Network





Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

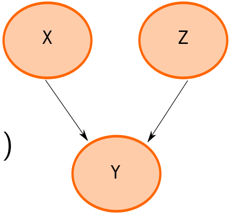
Deep Learning Strategies – Part III

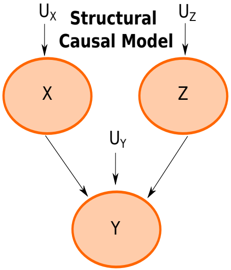
Causal Learning

Structural Causal Models (SCMs) – Example XOR

- $P(V) = P(X = x, Z = z, Y = y)$
- $P(V) = \prod_X P(X|Pa_X) = P(X = x|Pa_X) \cdot P(Z = z|Pa_Z) \cdot P(Y = y|Pa_Y)$
- $X = f_X(Pa_X, U_X) = U_x$
- $Z = f_Z(Pa_Z, U_z) = U_z$
- $Y = f_Y(Pa_Y, U_Y)$, with $U_Y = 0 \rightarrow Y = (X, Z)$
- $P(V) = P(x)P(z)P(y|x, z) = P(U_X = x)P(U_Z = z)1_y(x \oplus z)$
  with: $1_y(b) = 1$, if $y = b$, otherwise 0

Causal Network



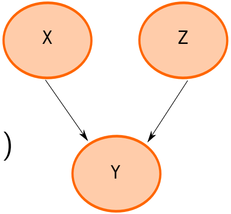


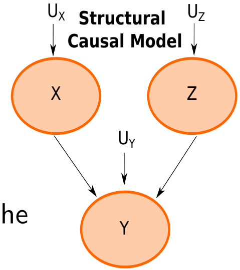
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR

- $P(V) = P(X = x, Z = z, Y = y)$
- $P(V) = \prod_X P(X|Pa_X) = P(X = x|Pa_X) \cdot P(Z = z|Pa_Z) \cdot P(Y = y|Pa_Y)$
- $X = f_X(Pa_X, U_X) = U_x$
- $Z = f_Z(Pa_Z, U_z) = U_z$
- $Y = f_Y(Pa_Y, U_Y)$, with $U_Y = 0 \rightarrow Y = (X, Z)$
- $P(V) = P(x)P(z)P(y|x, z) = P(U_X = x)P(U_Z = z)1_y(x \oplus z)$
  with: $1_y(b) = 1$, if $y = b$, otherwise 0
- Note: From an SCM it is possible to compute the joint distribution of the given variables!

Causal Network



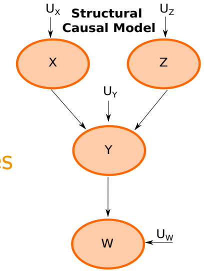


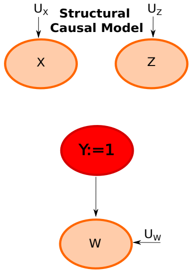
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Example XOR – Intervention

- Causal network turns into a Bayesian network if augmented with CPDs (prior probabilities, conditional probabilities) answering probability queries $P(Y|Z=z)$ for all $z \in Val(Z)$, $Z, Y \subseteq V$

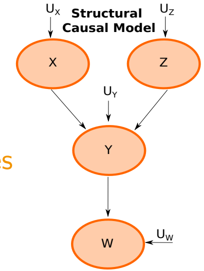


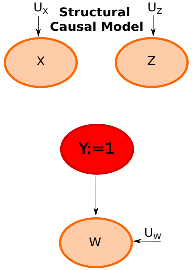


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Intervention

- Causal network turns into a Bayesian network if augmented with CPDs (prior probabilities, conditional probabilities) answering probability queries $P(Y|Z=z)$ for all $z \in Val(Z)$, $Z, Y \subseteq V$
- Conditional Probability Distribution (CPD): function that defines the probability of a random variable, given the values of its parent variables in a graph $P(X|Pa_X)$





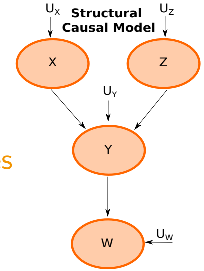
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

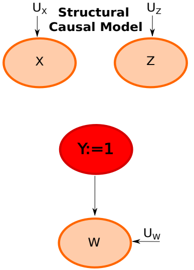
Structural Causal Models (SCMs) – Example XOR – Intervention

- Causal network turns into a Bayesian network if augmented with CPDs (prior probabilities, conditional probabilities) answering probability queries $P(Y|Z = z)$ for all $z \in Val(Z)$, $Z, Y \subseteq V$
- Conditional Probability Distribution (CPD): function that defines the probability of a random variable, given the values of its parent variables in a graph $P(X|Pa_X)$
- Intervention in SCMs (do-operation): setting $Y := 1$

  Before: $X := U_X,\ Z := U_Z,\ Y := X \oplus Z + U_Y,\ W := 2Y + U_W$

  After: $X := U_X,\ Z := U_Z,\ Y := 1,\ W := 2Y + U_W \rightarrow U_W = W - 2$





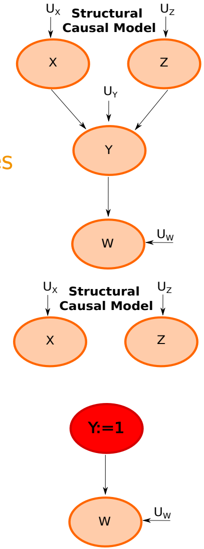
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Intervention

- Causal network turns into a Bayesian network if augmented with CPDs (prior probabilities, conditional probabilities) answering probability queries $P(Y|Z = z)$ for all $z \in Val(Z)$, $Z, Y \subseteq V$
- Conditional Probability Distribution (CPD): function that defines the probability of a random variable, given the values of its parent variables in a graph $P(X|Pa_X)$
- Intervention in SCMs (do-operation): setting $Y := 1$

  Before: $X := U_X, \; Z := U_Z, \; Y := X \oplus Z + U_Y, \; W := 2Y + U_W$

  After: $X := U_X, \; Z := U_Z, \; Y := 1, \; W := 2Y + U_W \rightarrow U_W = W - 2$
- Y does not depend on its endogenous/exogenous parents Pa_Y (all incoming links in the graph are removed!) $\rightarrow$ Perfect intervention!



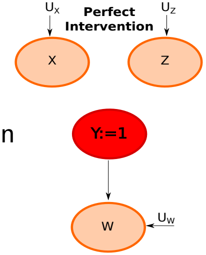
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

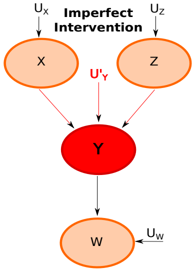
Deep Learning Strategies – Part III

Causal Learning

Structural Causal Models (SCMs) – Example XOR – Intervention

- In a perfect intervention the variable Y has a fixed/static value, based on its intervened value

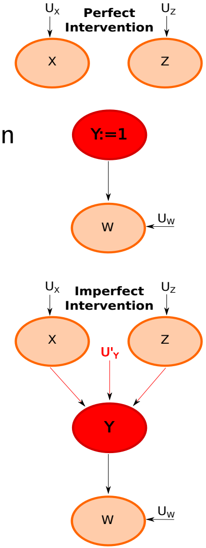




Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Intervention

- In a perfect intervention the variable Y has a fixed/static value, based on its intervened value
- However: there exist also an imperfect/parametric intervention, known as mechanism change (f' assignment) (still some remaining parents) $\rightarrow$ $Y := f(Pa_Y, U_Y) \rightarrow Y := f'(Pa_Y, U'_Y)$ (e.g. $Y := X - Z + U'_Y$)

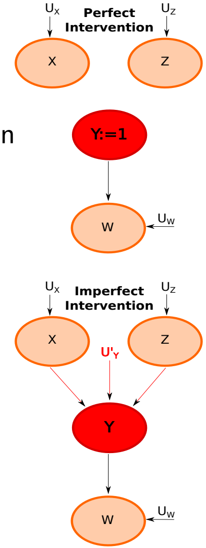


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Intervention

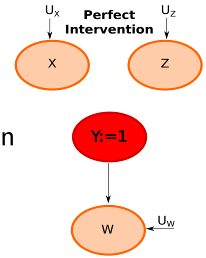
- In a perfect intervention the variable Y has a fixed/static value, based on its intervened value
- However: there exist also an imperfect/parametric intervention, known as mechanism change (f' assignment) (still some remaining parents) $\rightarrow$ $Y := f(Pa_Y, U_Y) \rightarrow Y := f'(Pa_Y, U'_Y)$ (e.g. $Y := X - Z + U'_Y$)
- Every intervention leads to a CPD change w.r.t. Y from:

  $Y : P(Y|Pa_Y)$ to $P^{\star}(Y|Pa_Y)$



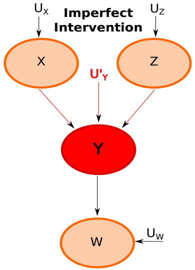
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Intervention



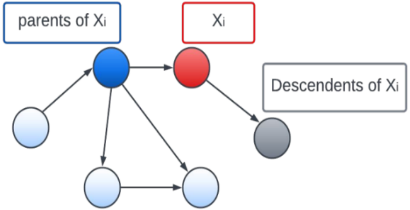
- In a perfect intervention the variable Y has a fixed/static value, based on its intervened value
- However: there exist also an imperfect/parametric intervention, known as mechanism change (f' assignment) (still some remaining parents) $\rightarrow$ $Y := f(Pa_Y, U_Y) \rightarrow Y := f'(Pa_Y, U'_Y)$ (e.g. $Y := X - Z + U'_Y$)
- Every intervention leads to a CPD change w.r.t. Y from:

  $Y : P(Y|Pa_Y)$ to $P^\star(Y|Pa_Y)$
- After every intervention there exist still a (new) SCM (causal network), with different dependencies and a intervention distribution $P(V\backslash Z, \mathrm{do}(Z))$ a intervention distribution over all the variables V



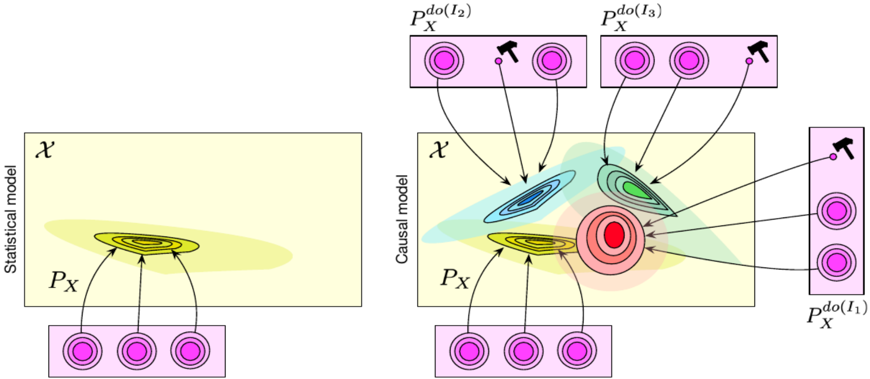
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Describing the causal mechanisms of a system.



- **Variables:** $\mathbf{X} = \{X_1, \ldots, X_N\}$
- **Noise:** $U = \{U_1, \ldots, U_N\}$
- **Causal parents of Xi** $X_{pa(i,G)}$
- **Structural Equations:** (Causal Mechanisms) $X_i = f_i\big(X_{pa(i,\mathbf{G})}, U_i\big) \quad \forall i \in \{1, \ldots N\}$

Source: Image from Nan Rosemary Ke, Stefan Bauer, "Causality and Deep Learning: Synergies, Challenges and the Future", Slide 22



- Statistical (Bayesian) Model: can only capture one probability distribution
- Causal Model: every intervention defines new joint distribution intervention

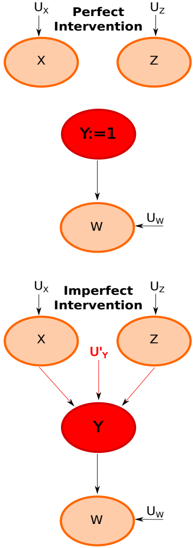
Source: FAU Erlangen-Nuremberg, Pattern Recognition Lab, K.Breininger, V. Christlein, Advanced Deep Learning – Interpretable/Causal Deep Learning

Source: Image from Nan Rosemary Ke, Stefan Bauer, "Causality and Deep Learning: Synergies, Challenges and the Future", Slide 22

Source: Schölkopf et al., "Toward Causal Representation Learning"

Structural Causal Models (SCMs) – Example XOR – Mutilated Network

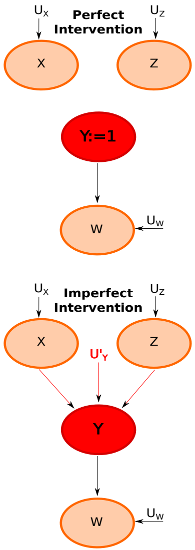
- Converting an intervention to an observation query (remove all do's)



Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Mutilated Network

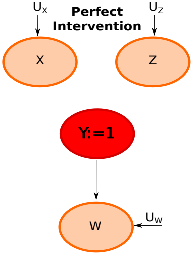
- Converting an intervention to an observation query (remove all do's)
- In case of the previous Bayesian Network, and for a perfect intervention query, $P(Y|\mathrm{do}(z), x)$ is transformed to $P(Y|z, x)$, by removing all incoming edges to the intervened nodes Z, and set a fixed intervention value $Z = z$

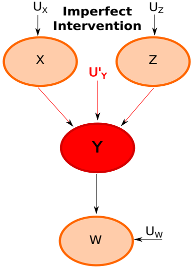


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Mutilated Network

- Converting an intervention to an observation query (remove all do's)
- In case of the previous Bayesian Network, and for a perfect intervention query, $P(Y|\text{do}(z), x)$ is transformed to $P(Y|z, x)$, by removing all incoming edges to the intervened nodes Z, and set a fixed intervention value $Z = z$
- Mutilated Network: $\mathcal{B}_{Z:=z}$, given a Bayesian Network $\mathcal{B}$, while removing all incoming links to the set of nodes Z and setting $Z = z$ (static value), $P(Z = z') = 1$ if $z' = z$, otherwise 0





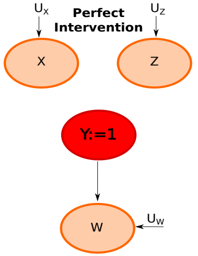
Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

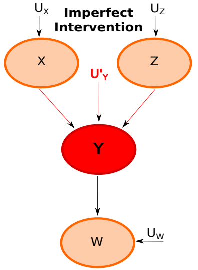
Deep Learning Strategies – Part III

Causal Learning

Structural Causal Models (SCMs) – Example XOR – Mutilated Network

- Converting an intervention to an observation query (remove all do's)
- In case of the previous Bayesian Network, and for a perfect intervention query, $P(Y|\text{do}(z), x)$ is transformed to $P(Y|z, x)$, by removing all incoming edges to the intervened nodes Z, and set a fixed intervention value $Z = z$
- Mutilated Network: $\mathcal{B}_{Z:=z}$, given a Bayesian Network $\mathcal{B}$, while removing all incoming links to the set of nodes Z and setting $Z = z$ (static value), $P(Z = z') = 1$ if $z' = z$, otherwise 0
- Answering intervention queries: $P(Y|\text{do}(z), x) = P_{\mathcal{B}_{Z:=z}}(Y|z, x)$

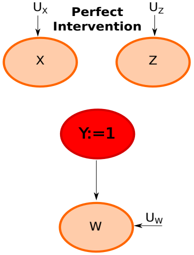


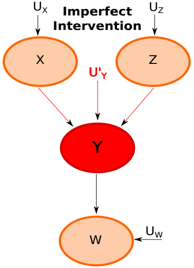


Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

Structural Causal Models (SCMs) – Example XOR – Mutilated Network

- Converting an intervention to an observation query (remove all do's)
- In case of the previous Bayesian Network, and for a perfect intervention query, $P(Y|\text{do}(z), x)$ is transformed to $P(Y|z, x)$, by removing all incoming edges to the intervened nodes Z, and set a fixed intervention value $Z = z$
- Mutilated Network: $\mathcal{B}_{Z:=z}$, given a Bayesian Network $\mathcal{B}$, while removing all incoming links to the set of nodes Z and setting $Z = z$ (static value), $P(Z = z') = 1$ if $z' = z$, otherwise 0
- Answering intervention queries: $P(Y|\text{do}(z), x) = P_{\mathcal{B}_{Z:=z}}(Y|z, x)$
- Modularity: conditional probability of a node is only affected if intervention is on its parents $P(Y|Pa_Y, \text{do}(z)) = P(Y|Pa_Y), \;\forall Y \notin Z$





Source: YouTube – easy learning – Causality 3: Defining causality: Structural causal models (SCM), Link

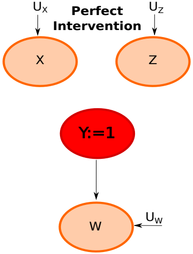
Deep Learning Strategies – Part III

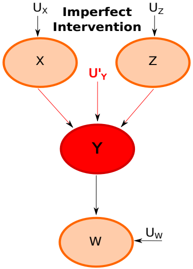
Causal Learning

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Example XOR – Causal (Bayesian) Model

- **Causal (Bayesian) Network $\mathcal{C}$ (CBN):** over a set of random variables V (joint distribution P), $\mathcal{C}$ is a Bayesian network $\mathcal{B}(\mathcal{G}, P_{\mathcal{C}})$, with $\mathcal{G}$ is an I-Map for P, over V, answering both probability queries:
  1. $P(Y|x) = P_{\mathcal{C}}(Y|x)$ and intervention queries:
  2. $P(Y|Pa_Y, \mathrm{do}(z)) = P_{\mathcal{C}_{Z:=z}}(Y|Pa_Y), \;\; \forall Y \notin Z, \;\; z \in \mathrm{Val}(Z)$



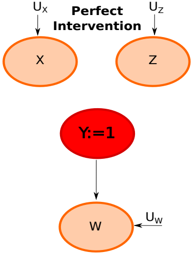


Source: YouTube – easy learning – Causality 4: Causal Bayesian Networks, Link

Deep Learning Strategies – Part III

Causal Learning

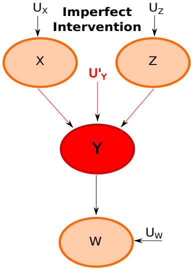
Structural Causal Models (SCMs) – Example XOR – Causal (Bayesian) Model



- Causal (Bayesian) Network $\mathcal{C}$ (CBN): over a set of random variables V (joint distribution P), $\mathcal{C}$ is a Bayesian network $\mathcal{B}(\mathcal{G}, P_{\mathcal{C}})$, with $\mathcal{G}$ is an I-Map for P, over V, answering both probability queries:
  1. $P(Y|x) = P_{\mathcal{C}}(Y|x)$ and intervention queries:
  2. $P(Y|Pa_Y, \text{do}(z)) = P_{\mathcal{C}_{Z:=z}}(Y|Pa_Y), \;\; \forall Y \notin Z, \;\; z \in \text{Val}(Z)$

- Recap: Causal Hierarchy

1. Association (Seeing!) – Probabilistic query $P(Y|X)$ (e.g. probability of developing cancer given that the individual smokes?) – BN, CBN, SCM
2. Intervention (Doing!) – Invervention query $P(Y|\text{do}(X))$ (e.g. What is the prob. of developing cancer if individuals stop smoking?) – CBN, SCM
3. Counterfactual (Imagining!) – Counterfactual query $P(Y_{X=1}|X = 0)$ (e.g. What would have happpened if individuals had quitted smoking?) – SCM



Source: YouTube – easy learning – Causality 4: Causal Bayesian Networks, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Bayesian Network (BN) is able to answer probabilistic queries (associations), while Causal Bayesian Networks (CBN) & SCM can also handle intervention queries

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Counterfactual Queries

- Bayesian Network (BN) is able to answer probabilistic queries (associations), while Causal Bayesian Networks (CBN) & SCM can also handle intervention queries
- What about the last topic, addressing counterfactual queries?

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Counterfactual Queries

- Bayesian Network (BN) is able to answer probabilistic queries (associations), while Causal Bayesian Networks (CBN) & SCM can also handle intervention queries
- What about the last topic, addressing counterfactual queries?
- Example: There was a fork in the road, where a decision must be made – take the freeway ($X = 1$) or go on a surface street ($X = 0$). The person took the surface street with a lot of traffic and arrived one hour later ($Y = 1$). He asked hisself – If he had taken the freeway, he would have gotton home earlier!!!

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Bayesian Network (BN) is able to answer probabilistic queries (associations), while Causal Bayesian Networks (CBN) & SCM can also handle intervention queries
- What about the last topic, addressing counterfactual queries?
- Example: There was a fork in the road, where a decision must be made – take the freeway ($X = 1$) or go on a surface street ($X = 0$). The person took the surface street with a lot of traffic and arrived one hour later ($Y = 1$). He asked hisself – If he had taken the freeway, he would have gotton home earlier!!!
- Such an unrealized situation from the past, wrapped in an "if-statement" is known as counterfactual $\rightarrow$ How to model a counterfactual query/situation?

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Counterfactual Queries

- Let Y be the arriving time ($Y = 1$ for being late, $Y = 1$ for being in time)

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Counterfactual Queries

- Let Y be the arriving time ($Y = 1$ for being late, $Y = 1$ for being in time)
- Causal inference: prediction of Y from the past, by intervening on the road (change from surface road $X = 0$ to freeway $X = 1$): $P(Y|\text{do}(X = 1))$

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Let Y be the arriving time ($Y = 1$ for being late, $Y = 1$ for being in time)
- Causal inference: prediction of Y from the past, by intervening on the road (change from surface road $X = 0$ to freeway $X = 1$): $P(Y|\text{do}(X = 1))$
- Counterfactual problem: asking about Y, if we had forced $X := 1$ in the past observed event with $X = 0, Y = 1 \rightarrow$ Clash regarding Y: $P(Y|\text{do}(X = 1), Y = 1, X = 0)$

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

OTH Amberg-Weiden

Structural Causal Models (SCMs) – Counterfactual Queries

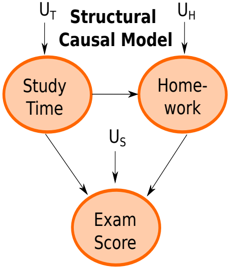
- Let Y be the arriving time ($Y = 1$ for being late, $Y = 1$ for being in time)
- Causal inference: prediction of Y from the past, by intervening on the road (change from surface road $X = 0$ to freeway $X = 1$): $P(Y|\text{do}(X = 1))$
- Counterfactual problem: asking about Y, if we had forced $X := 1$ in the past observed event with $X = 0, Y = 1 \rightarrow$ Clash regarding Y: $P(Y|\text{do}(X = 1), Y = 1, X = 0)$
- Counterfactual inference: reformulate as: $P(Y_{X=1}|X = 0, Y = 1)$

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Example Counterfactual Event:
  - $T := U_T, \quad H := 0.5T + U_H, \quad S := 0.7T + 0.4H + U_S$

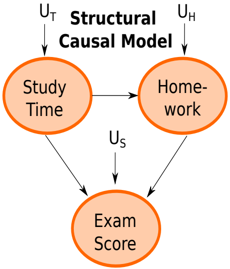
    with: T as the students studying time, H as the number of completed homeworks, S student's exam score, U_T student's free time at home, U_H student's intelligence, U_S student's anxiety



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

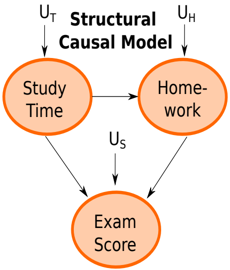
- Example Counterfactual Event:
  - $T := U_T, \quad H := 0.5T + U_H, \quad S := 0.7T + 0.4H + U_S$
    with: T as the students studying time, H as the number of completed homeworks, S student's exam score, U_T student's free time at home, U_H student's intelligence, U_S student's anxiety
  - Let $T = 0.5$, $H = 1$, and $S = 1.5$ for a particular student, resulting in:
    $U_T := T = 0.5$, $U_H = H - 0.5T = 1 - 0.5 \cdot 0.5 = 0.75$,
    $U_S = S - 0.7T - 0.4H = 1.5 - 0.7 \cdot 0.5 - 0.4 \cdot 1 = 0.75$



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

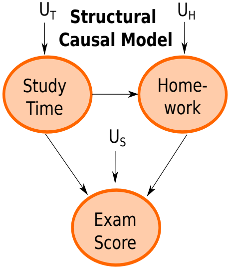
- Example Counterfactual Event:
  - $T := U_T, \quad H := 0.5T + U_H, \quad S := 0.7T + 0.4H + U_S$
    with: T as the students studying time, H as the number of completed homeworks, S student's exam score, U_T student's free time at home, U_H student's intelligence, U_S student's anxiety
  - Let $T = 0.5$, $H = 1$, and $S = 1.5$ for a particular student, resulting in:
    $U_T := T = 0.5$, $U_H = H - 0.5T = 1 - 0.5 \cdot 0.5 = 0.75$,
    $U_S = S - 0.7T - 0.4H = 1.5 - 0.7 \cdot 0.5 - 0.4 \cdot 1 = 0.75$
  - Counterfactual problem: what would the student's score have been if he had doubled his done homework?



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

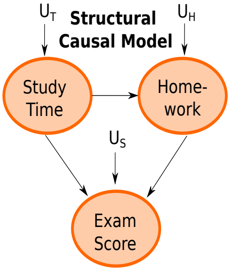
- Example Counterfactual Event:
  - $T := U_T, \quad H := 0.5T + U_H, \quad S := 0.7T + 0.4H + U_S$
    with: T as the students studying time, H as the number of completed homeworks, S student's exam score, U_T student's free time at home, U_H student's intelligence, U_S student's anxiety
  - Let $T = 0.5$, $H = 1$, and $S = 1.5$ for a particular student, resulting in:
    $U_T := T = 0.5$, $U_H = H - 0.5T = 1 - 0.5 \cdot 0.5 = 0.75$,
    $U_S = S - 0.7T - 0.4H = 1.5 - 0.7 \cdot 0.5 - 0.4 \cdot 1 = 0.75$
  - Counterfactual problem: what would the student's score have been if he had doubled his done homework?
  - Intervention: do-operator to $P(S|Pa_S, U_S) = P(S|T, \mathrm{do}(H := 2), U_S)$



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

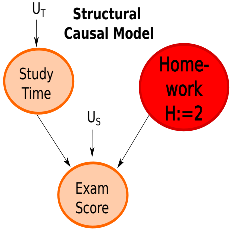
- Example Counterfactual Event:
  - $T := U_T, \quad H := 0.5T + U_H, \quad S := 0.7T + 0.4H + U_S$
    with: T as the students studying time, H as the number of completed homeworks, S student's exam score, U_T student's free time at home, U_H student's intelligence, U_S student's anxiety
  - Let $T = 0.5$, $H = 1$, and $S = 1.5$ for a particular student, resulting in:
    $U_T := T = 0.5$, $U_H = H - 0.5T = 1 - 0.5 \cdot 0.5 = 0.75$,
    $U_S = S - 0.7T - 0.4H = 1.5 - 0.7 \cdot 0.5 - 0.4 \cdot 1 = 0.75$
  - Counterfactual problem: what would the student's score have been if he had doubled his done homework?
  - Intervention: do-operator to $P(S|Pa_S, U_S) = P(S|T, \mathrm{do}(H := 2), U_S)$
  - Counterfaction: $P(S_{H=2}|T = 0.5, H = 1, S = 1.5)$



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

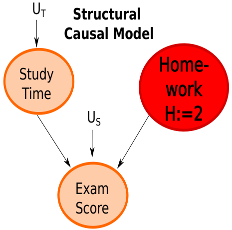
- Example Counterfactual Event
  - Create mutilated graph, by removing all incoming links to the invervened node $Z = H$, with $P(H = 2) = 1$



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

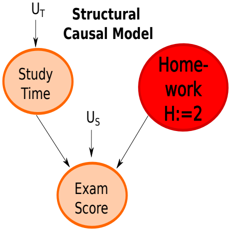
- Example Counterfactual Event
  - Create mutilated graph, by removing all incoming links to the invervened node $Z = H$, with $P(H = 2) = 1$
  - Compute S in the mutilated graph, by $S := 0.7T + 0.4H + U_S$ with: $S_{H=2}(T = 0.5, U_S = 0.75) = 0.7 \cdot 0.5 + 0.4 \cdot 2 + 0.75 = 1.9$



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

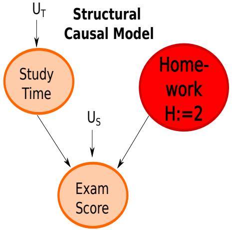
- Example Counterfactual Event
  - Create mutilated graph, by removing all incoming links to the invervened node $Z = H$, with $P(H = 2) = 1$
  - Compute S in the mutilated graph, by $S := 0.7T + 0.4H + U_S$ with: $S_{H=2}(T = 0.5, U_S = 0.75) = 0.7 \cdot 0.5 + 0.4 \cdot 2 + 0.75 = 1.9$
- Abduction-Action-Prediction
  1. Abduction – use evidence $E = e$ in order to determine all the values of U



Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Example Counterfactual Event
  - Create mutilated graph, by removing all incoming links to the invervened node $Z = H$, with $P(H = 2) = 1$
  - Compute S in the mutilated graph, by $S := 0.7T + 0.4H + U_S$ with: $S_{H=2}(T = 0.5, U_S = 0.75) = 0.7 \cdot 0.5 + 0.4 \cdot 2 + 0.75 = 1.9$

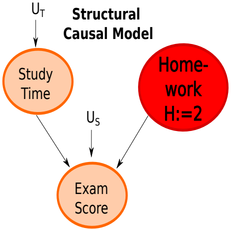


- Abduction-Action-Prediction
  1. Abduction – use evidence $E = e$ in order to determine all the values of U
  2. Action – Removing structural equations for variables X and replace it with the appropriate functions $X = x$ (Mutilated Graph!)

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Structural Causal Models (SCMs) – Counterfactual Queries

- Example Counterfactual Event
  - Create mutilated graph, by removing all incoming links to the invervened node $Z = H$, with $P(H = 2) = 1$
  - Compute S in the mutilated graph, by $S := 0.7T + 0.4H + U_S$ with: $S_{H=2}(T = 0.5, U_S = 0.75) = 0.7 \cdot 0.5 + 0.4 \cdot 2 + 0.75 = 1.9$



- Abduction-Action-Prediction
  1. Abduction – use evidence $E = e$ in order to determine all the values of U
  2. Action – Removing structural equations for variables X and replace it with the appropriate functions $X = x$ (Mutilated Graph!)
  3. Prediction – Use the mutilated model and the value of U to compute the value of Y, the consequence of the counterfactual

Source: YouTube – easy learning – Causality 6: Counterfactual Queries, Link

Summary – Difference BN, CBN and SCM

- Bayesian Network (BN)

1. BN is a directed acyclic graph (DAG) that represents probabilistic relationships among variables, while each random variable (node) connection represents a conditional dependence $\rightarrow$ No causation, only correlations, and conditional independencies

Summary – Difference BN, CBN and SCM

- Bayesian Network (BN)

1. BN is a directed acyclic graph (DAG) that represents probabilistic relationships among variables, while each random variable (node) connection represents a conditional dependence → No causation, only correlations, and conditional independencies
2. Example: Bayesian Network, with nodes: Flu → Fever, Flu → Cough, while the model/graph shows the association between Flu and Fever, as well as Flu and Cough, however, no information that Flue causes both

OTH Amberg-Weiden

Summary – Difference BN, CBN and SCM

- Causal Bayesian Network (CBN)

1. CBN is a BN with an additional assumption – the edges represent causal relationships – explicitly encoding causation, not just correlation or conditional independence

OTH Amberg-Weiden

Summary – Difference BN, CBN and SCM

- Causal Bayesian Network (CBN)

1. CBN is a BN with an additional assumption – the edges represent causal relationships – explicitly encoding causation, not just correlation or conditional independence
2. Each edge is interpreted as causal link, describing a change in one variable directly causes a change in the connected variable

Summary – Difference BN, CBN and SCM

- Causal Bayesian Network (CBN)

1. CBN is a BN with an additional assumption – the edges represent causal relationships – explicitly encoding causation, not just correlation or conditional independence
2. Each edge is interpreted as causal link, describing a change in one variable directly causes a change in the connected variable
3. CBN allows to perform interventions (e.g., forcing a variable to a particular value) and predict how this intervention will impact other variables

OTH Amberg-Weiden

Summary – Difference BN, CBN and SCM

- Causal Bayesian Network (CBN)

1. CBN is a BN with an additional assumption – the edges represent causal relationships – explicitly encoding causation, not just correlation or conditional independence
2. Each edge is interpreted as causal link, describing a change in one variable directly causes a change in the connected variable
3. CBN allows to perform interventions (e.g., forcing a variable to a particular value) and predict how this intervention will impact other variables
4. CBNs are used for causal inference, with predictions not just about associations, but also the effects of interventions

Summary – Difference BN, CBN and SCM

- Causal Bayesian Network (CBN)

1. CBN is a BN with an additional assumption – the edges represent causal relationships – explicitly encoding causation, not just correlation or conditional independence
2. Each edge is interpreted as causal link, describing a change in one variable directly causes a change in the connected variable
3. CBN allows to perform interventions (e.g., forcing a variable to a particular value) and predict how this intervention will impact other variables
4. CBNs are used for causal inference, with predictions not just about associations, but also the effects of interventions
5. Example: same graph w.r.t. causation, show that Flu causes Fever and Cough, while intervening Flu causes an impact with resepct to Fever and Cough

Summary – Difference BN, CBN and SCM

- Structural Causal Model (SCM)

1. SCM is a causal model providing a more formal framework than a CBN, using a set of structural equations that specify how each variable is generated as a function of its causal parents and some random "noise" term, along with a causal DAG structure

Summary – Difference BN, CBN and SCM

- Structural Causal Model (SCM)

1. SCM is a causal model providing a more formal framework than a CBN, using a set of structural equations that specify how each variable is generated as a function of its causal parents and some random "noise" term, along with a causal DAG structure
2. SCMs allow for both causal inference and formal intervention analysis, while being used for counterfactual reasoning, interventions, and predicting causal effects

Summary – Difference BN, CBN and SCM

- Structural Causal Model (SCM)

1. SCM is a causal model providing a more formal framework than a CBN, using a set of structural equations that specify how each variable is generated as a function of its causal parents and some random "noise" term, along with a causal DAG structure
2. SCMs allow for both causal inference and formal intervention analysis, while being used for counterfactual reasoning, interventions, and predicting causal effects
3. SCM would not only represent Flu causing Fever & Cough, but also provide equations – Fever$=f(\text{Flu}, U_F)$, Cough$=g(\text{Flu}, U_C)$ – specifying how Flu influences Fever & Cough under different conditions

Further Questions?

https://www.oth-aw.de/hochschule/ueber-uns/personen/bergler-christian/

Source: https://emekaboris.medium.com/the-intuition-behind-100-days-of-data-science-code-c98402cdc92c