

شبکه عصبی همیلتونی

فاطمه حافظیانزاده^۱

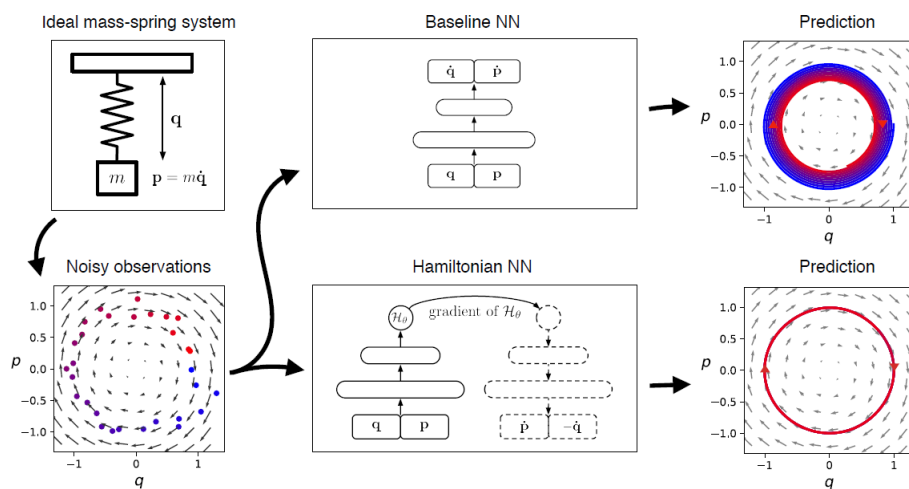
استاد راهنما: دکتر احسان ندایی اسکوئی^۲

دانشکده فیزیک، دانشگاه تحصیلات تکمیلی علوم پایه زنجان، صندوق پستی ۴۵۱۹۵-۱۱۵۹، زنجان، ایران

یکی از مسائلی که در شبکه‌های عصبی چالش برانگیز است، یادگیری شبکه‌های عصبی است که بر اساس قوانین ساده‌ی فیزیکی هستند، در این پژوهش به بررسی تشکیل و تربیت شبکه‌ای عصبی بر اساس مکانیک هامیلتونی پرداخته شده‌است و به مقایسه‌ی شبکه‌هایی که از قانون پایستگی انرژی تبعیت می‌کنند با شبکه‌هایی که بدون توجه به مکانیک هامیلتونی و پایستگی انرژی تربیت شده‌اند، پرداخته شده‌است.

۱ مقدمه

شبکه‌های عصبی توانایی قابل توجهی در یادگیری داده‌ها و به دنبال آن پیش‌بینی داده‌ها دارند. شبکه‌های عصبی در طبقه‌بندی تصاویر، یادگیری تقویتی و مهارت رباتیک سرآمد هستند. در تمامی این موارد قوانین فیزیکی ثابت و یکسان است. این موضوع باعث شده تا محققان شبکه‌های عصبی به سمت فیزیک آمده و سعی کنند تا قوانین فیزیکی را در قسمت یادگیری شبکه‌های عصبی وارد کنند. در این پژوهش نیز سعی شده شبکه‌ی عصبی برای پیش‌بینی حرکت جرم و فنر به گونه‌ای تربیت شود که از قانون پایستگی انرژی تبعیت کند.



شکل ۱: شکل شماتیک به منظور مقایسه‌ی دو مدل طراحی شده جرم و فنر در این پژوهش برای یادگیری شبکه عصبی همیلتونی (HNN) و شبکه عصبی پایه (baseline) با وارد کردن اختلال در داده‌های آزمایش مدل baseline تربیت شده در پیش‌بینی مسیر دچار اختلال می‌شود در حالی که مدل HNN به خوبی مسیر را پیش‌بینی می‌کند.

۲ تئوری

۱. پیش‌بینی دینامیک: یک مدل فیزیکی خوب، مدلی است که بتواند مسیر حرکت سیستم را پیش‌بینی کند. ساده‌ترین مدلی که می‌توان مطرح نمود مدلی است که مکان و سرعت سیستم را در یک گام زمانی بعد پیش‌بینی کند. نام این مدل در این پژوهش مدل پایه^۳ قرار

¹Email: fatemehhafezi@iasbs.ac.ir

²Email: nedaace@iasbs.ac.ir

³baseline model

گرفته شده است. ولی این مدل طراحی شده این مشکلات را داشت: ۱. سخت بودن تعریف کردن گام زمانی ۲. با وارد کردن یک اختلال جزئی در مکان و سرعت‌های اولیه به دلیل این که این مدل از قانون پایستگی انرژی پیروی نمی کند، سیستم از مسیر منحرف می شود. که هر دو این مشکل‌ها در مدل شبکه عصبی همیلتونی^۴ حل شدند.

۲. مکانیک همیلتونی: ویلیام همیلتون^۵ مکانیک همیلتونی را در قرن ۱۹ به عنوان یک فرمول ریاضی وارد مکانیک کلاسیک کرد. بعد از آن دانشمندان آن را وارد حوزه‌های مختلف فیزیک، از ترمودینامیک گرفته تا نظریه میدان‌های کوانتومی کردند. در مکانیک همیلتونی به هر ذره دو کمیت برداری مکان و تکانه (q, p) و یک کمیت نرده ای با عنوان همیلتونی H نسبت داده می شود. که روابط معادله حرکت به این گونه تعریف می شوند:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial q}$$

۳. شبکه‌ی عصبی همیلتونی: در این پژوهش، به صورت غیرنظارتی کمیتی نرده‌ای (انرژی گونه) محاسبه می شود و قبل از محاسبه‌ی تابع زیان^۶ به روش گراف-گرادیان از آن کمیت انرژی گونه مشتق گرفته می شود و مقادیر مشتقات مکان و تکانه محاسبه می شوند. سپس تابع زیان زیر بهینه می شود:

$$L_{HNN} = \left\| \frac{\partial H}{\partial p} - \frac{\partial q}{\partial t} \right\|_2 + \left\| \frac{\partial H}{\partial q} + \frac{\partial p}{\partial t} \right\|_2$$

با انجام مراحل گفته شده، در مدل شبکه عصبی همیلتونی، انرژی کل پایسته باقی می ماند یعنی شبکه‌ی عصبی طراحی می شود که در آن انرژی بر اساس مکانیک همیلتونی پایسته باقی می ماند.

۳ شرح مسئله

مدل را با نرخ یادگیری^{۳-۱۰} و بهینه ساز ADAM تربیت کردیم. برای همه‌ی مجموعه داده‌های دو شبکه عصبی کاملاً متصل را تربیت کردیم: مدل پایه ای: در این مدل به عنوان ورودی (q, p) گرفته می شود و به عنوان خروجی به صورت مستقیم (\dot{q}, \dot{p}) دریافت می شود. مدل شبکه عصبی همیلتونی: (q, p) به عنوان ورودی گرفته می شود و سپس در مرحله ی جلو^۷ کمیت اسکالری تحت عنوان همیلتونی محاسبه می شود و در مرحله ی روبه عقب^۸ طبق معادلات حرکت همیلتونی از آن مشتق گرفته می شود و در نهایت به عنوان خروجی $(\dot{p}, -\dot{q})$ دریافت می شود (شکل ۱). در هر دو مدل ها سه لایه ی عصبی با ۲۰۰ واحد پنهان با تابع فعالیت \tanh قرار گرفته شده است. برای محاسبه ی دینامیک سیستم از روش $Runge - Kutta$ استفاده شد.

۴ نتایج

مدل HNN که براساس پایستگی انرژی طراحی شد؛ مسیر حرکت جرم و فنر را به طور دقیق پیش بینی می کند. حتی با وارد کردن اختلال در داده‌های آموزشی شبکه به گونه‌ای تربیت می شود تا انرژی پایسته باقی بماند. این درحالی است که مدل baseline مسیر حرکت را به خوبی پیش بینی نمی کند و انرژی در آن پایسته باقی نمی ماند (شکل ۲).

مراجع

Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. In Advances in Neural Information Processing Systems, pp. 15353–15363, 2019.

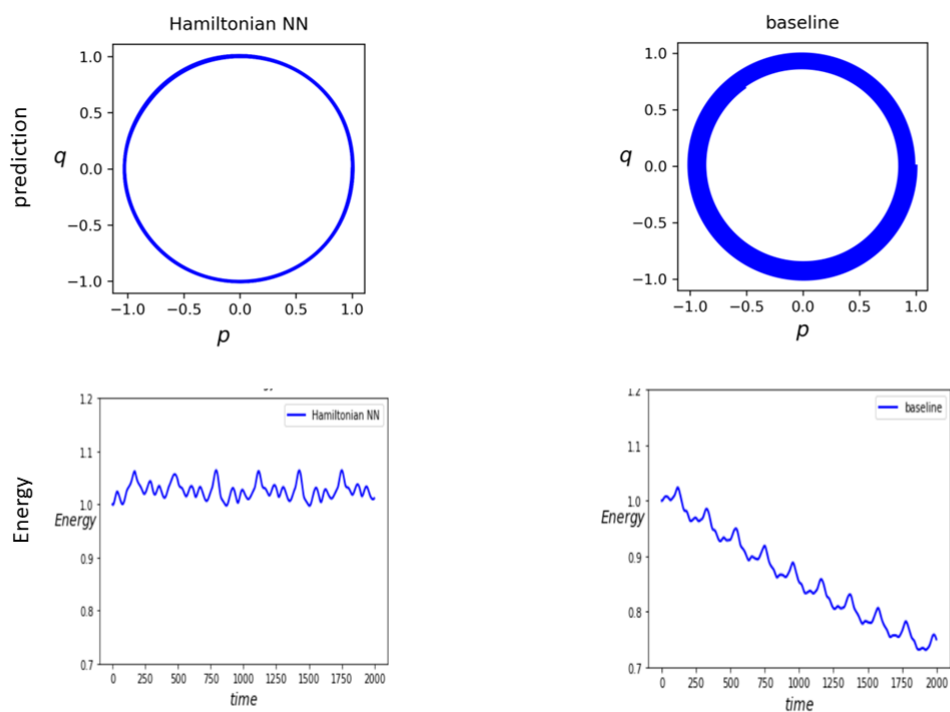
⁴Hamiltonian Neural Network model

⁵William Hamilton

⁶loss function

⁷forward

⁸backward



شکل ۲: شکل شماتیک مقایسه‌ی دو مدل HNN و baseline. ستون سمت راست؛ نتایج مربوط به مدل HNN و ستون سمت چپ نتایج مربوط به مدل baseline می‌باشد. ردیف اول نمودارهای پیش‌بینی مسیر و ردیف دوم نمودارهای مربوط به انرژی بر حسب گام‌های زمانی را نشان می‌دهد.